Coherent Spectroscopy at the Diffraction Limit

by

Eric W. Martin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Applied Physics)
in The University of Michigan
2018

Doctoral Committee:

Professor Steven T. Cundiff, Chair
Professor Mackillo Kira
Professor Roberto Merlin
Professor Theordore Norris
Professor Jennifer Ogilvie

Eric W. Martin

ewmartin@umich.edu

ORCID iD: 0000-0002-3707-3233

ACKNOWLEDGEMENTS

All of the work in this thesis would not have been possible without such a supportive and tight-knit lab. I have gained ideas and valuable insight from long conversations with many of my lab mates, and I am thankful for all of them. In particular I would like to acknowledge Rohan Singh and Chris Smallwood for being especially available to discuss experimental interpretations and theory. The experiments I have done have relied on a lot of the work of Travis Autry and Gaël Nardin, and I was very lucky to work next to them in Colorado. I am thankful to Albert Liu for reading over this thesis and providing feedback. I would also like to acknowledge the now well-bonded group that moved with me from Colorado to Michigan: Bachana Lomsadze, Diogo Almeida, Takeshi Suzuki, Kai Wang, Rohan Singh, and Chris Smallwood.

For their love and support, I thank my family and all my friends. Mom, Dad, and my brother Dillon have encouraged me and enabled my pursuit of science. I thank my fiancé Candice for putting up with long days in the lab and for her support over the last three years.

Chapters IV and V have been based on theory and insight by Prof. Mackillo Kira. I have benefited from many discussions with Mack, and I am thankful for the time he has given me. His student Markus Borsch has also worked hard to develop the inhomogeneous semiconductor Bloch equations that he is using to model the experimental results presented in Chapter V.

Chapters VII and VIII present measurements of samples created and characterized (using linear spectroscopy and atomic-force microscopy) in Prof. Hui Deng's group

by Jason Horng, Michael-Henr Wentzel, and Eunice Paik. Jason Horng, a post-doc in Prof. Hui Deng's group, has been exceptionally great to work with on these projects. He devised and developed most of the samples, initiated this collaboration by realizing how our experiments could be used to advance measurements of his samples, and has a strong intuition for the theory. Specifically he is also responsible for realizing many of the findings in Chapter VIII including the critical coupling condition and the analogue to coherent perfect absorption. The multidimensional coherent spectroscopy (MDCS) measurements for these chapters have been made with the help of Hanna G. Ruth. Hanna has also assisted with the analysis and interpretation of the results.

Chapter IX presents MDCS measurements on zero-dimensional polariton devices developed in Hui Deng's lab here at the University of Michigan. The devices were developed by Bo Zhang. I have directly worked with Jiaqi Hu and Zhaorong (Joey) Wang who have helped me with finding the directions to go in these studies and with finding the relevance of the MDCS results to the polariton community.

My advisor, Prof. Steve Cundiff, has of course played a tremendous role in all of this work. I would like to acknowledge him for giving me the perfect amount of freedom, support, and enthusiasm to do these projects. I have been allowed to pursue many ideas and topics, and his insight and guidance throughout the years have kept me on track.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

LIST OF ABBREVIATIONS

AOM acousto-optic modulator

ASE amplified stimulated emission

CVD chemical vapor deposition

CW continuous-wave

DBR distributed Bragg reflector

ECDL external-cavity diode laser

EID excitation-induced dephasing

EIS excitation-induced shift

FPGA field-programmable gate array

FrFT fractional Fourier transforms

FTIR Fourier-transform infrared

FWM four-wave mixing

GaAs gallium arsenide

hBN hexagonal boron nitride

LO local oscillator

LP lower polariton

MDCS multidimensional coherent spectroscopy

NA numerical aperture

OBE optical Bloch equation

PL photoluminescence

SNR signal-to-noise ratio

SWG sub-wavelength grating

TMDC transition metal dichalcogenide

TE transverse-electric

TM transverse-magnetic

QD quantum dot

QW quantum well

ABSTRACT

Semiconductor nanostructures have spatial inhomogeneity resulting from disorder, device design, and even experimental conditions. Inhomogeneity on one-micron length scales can be accessed with a diffraction-limited spot. Optical spectroscopies with the ability to measure on this length scale are capable of probing small samples and short-range effects that are otherwise inaccessible. In this thesis we develop a set of coherent spectroscopies based on heterodyne detection that enable measurement of ultrafast (sub-picosecond) dynamics in a diffraction-limited spot. Coherent spectroscopy techniques are known for their ability to probe many-body effects and microscopic inhomogeneity. We exploit these capabilities in several different experiments to measure effects at small length scales with extremely high sensitivity.

The second-order coherence of an optical excitation has known and predicted affects on nonlinear semiconductor responses. Here we develop an experiment to compare the optical excitation of gallium arsenide quantum wells using a truly thermal light source (a luminescent diode) and a coherent light source (laser). Through careful experimentation we measure no difference in these unique excitations, and we present that the absence of an effect arises in part because the measured nonlinear response results from averaging over multiple temporal modes of the continuous thermal source.

Experiments here demonstrate the ability to measure many-body effects on extremely short length scales, even less than the spot size. We measure nanometer dynamics of an excitation spot that result from the interplay between a tightly-focused ultrafast-laser spot and the collective excitation the laser induces in a semiconductor.

Specifically we use spatially-resolved transient absorption to measure that a small spot of resonantly excited excitons in 10 quantum well sample decreases in size for several picoseconds before the diffusion of excitons dominates the spot-size dynamics.

In a single narrow quantum well containing interfacial quantum dot states we use double-quantum multidimensional spectroscopy to measure coupling between nominally isolated localized (quantum dot) states. The coupling is only present after weak excitation of delocalized (quantum well) states in the same crystal. This experiment also demonstrates the ability of diffraction limited coherent techniques to probe length scales beyond the diffraction limit.

We use multidimensional coherent spectroscopy (MDCS) to compare monolayer molybdenum diselenide samples mounted directly on sapphire versus encapsulated in hexagonal noron nitride. We measure the homogeneous linewidths of various samples and demonstrate that the linewidth of the encapsulated samples are nearly radiatively limited at 0.26 ± 0.02 meV. The linewidth is much narrower than is expected based on the theoretical prediction for a perfect sample and the integrated absorption assuming a homogeneous sample. We attribute the narrow linewidth to localization and discuss inhomogeneity applied to highly coherent excitations.

Using MDCS we measure coupling between quantum-confined polariton states. The states are spatially localized, so we are able to measure the dependence of the signal on the spatial position of the exciation spots. We also measure relaxation from the exciton into the polariton states that indicate how energy is transferred when these systems are pumped as lasers.

Finally we discuss and demonstrate nonlocal four-wave mixing. We present this as a phenomenon that can be measured in materials with a high coherence length, and we present the technique as a way to determine a sample's coherence length. The fractional-Fourier transform is introduced for considering nonlinear responses in system with finite spatial extents.

CHAPTER I

Introduction

Nonlinear optical spectroscopy has proved incredibly useful for directly measuring many-body and correlation effects in semiconductors and semiconductor nanostructures [7, 8]. Consideration of these correlation effects has been essential in the theoretical treatment of semiconductors for correctly determining the properties of semiconductor laser diodes [7, 9, 10]. Laser diodes have transformed society with applications in laser printers, telecommunications, laser-based range sensors (e.g., Lidar), medical devices, and laser-light shows. Now novel materials, including atom-thick stackable monolayers, and semiconductor devices are being developed to extend the photonic and electronic applications of semiconductors. A careful consideration of the many-body effects in these new materials and nanostructures is more important than ever.

We study some of the many-body effects in semiconductors by using lasers to excite excitons, quasiparticle excitations in a semiconductor that result from Coulomb interactions between quasiparticle electrons and quasiparticle holes in a semiconductor crystal. In direct bandgap semiconductors excitons are relatively long-lived states that can be coherently excited by resonant light. By resonantly exciting excitons, it is possible to probe excitation induced shifting and broadening of the exciton resonance without creating free electrons and holes. Resonant nonlinear spectroscopies have thus been especially useful for studying the theory of homogeneous excitations

in semiconductors [8, 11].

Inhomogeneous excitations with resonant spectroscopies have been largely unexplored. Measuring inhomogeneous excitations requires the use of a small resonant excitation, and the smallest spot we can create by purely optical means is a diffraction-limited spot. There are three major features of a diffraction-limited spot that we have exploited to generate the results in this thesis.

1. We achieve high intensities with low power light sources.

2. We use the high spatial inhomogeneity of the focused light to apply a gradient force on particles.

3. We select regions of highly inhomogeneous samples.

In the following chapters I present exciting results that have been enabled by these features, but of course there have been challenges to overcome in order to do nonlinear spectroscopy in a diffraction-limited spot. One major challenge for any diffraction-limited nonlinear spectroscopy is that it prevents the use of wave-vector selection to isolate the nonlinear signal. This has often been overcome by performing two-color experiments like non-degenerate pump-probe and Raman spectroscopy that spectrally isolate the nonlinear signal. Here we use fully degenerate resonant spectroscopies to explore the coherent effects in a diffraction-limited spot. These methods are discussed throughout the thesis.

Another challenge has been the sensitivity of measurements on the sample position and focusing objectives. Though these solutions are not as completely highlighted in the thesis, no small amount of my time has been devoted to minimizing sample drift in a cryostat and speeding up scan durations.

1.1 High Nonlinearity with Low Power

In Chapter IV I discuss my first major project in the lab. The goal of this project was to excite a nonlinear response in a semiconductor with nonclassical light. Specifically, the nonlinear responses of coherent laser light and thermal light having equal average intensities should differ due to the high intensity fluctuations of the thermal light. The problem is that it is very difficult to generate a bright thermal light source capable of driving a nonlinear response in a material. It really is no coincidence that the first nonlinear optical process was discovered in 1961, one year after the invention of the laser [12]. However, we measured a sample with a high oscillator strength and we used a tight focus with a relatively high power luminescent light source to induce a nonlinear response. A careful experiment revealed that the excitation by a thermal source was not quite the same as excitation by a thermal quantum optical state. We present our results and an outlook on potential methods for moving this project forward.

Though we have used a diffraction-limited spot to enable the otherwise impossible access to the nonlinear regime with a thermal light source, the ability to induce a nonlinear response with weak intensity light is a matter of practicality as well. Many nonlinear spectroscopy techniques in research labs are performed with expensive amplified laser systems that output extremely high peak intensities. These lasers are essential for inducing measurable nonlinearity in low density samples with low dipole moments, but this is partly due to the fact that the techniques utilize wave vector selection to isolate the nonlinear signal from the excition beams. These non-collinear geometries have spot sizes that are fundamentally limited to be greater than 30 μm. We have managed to measure the same signals with excitation spot sizes that are almost 100 times smaller than those required by the state-of-the-art non-collinear alternatives. The spot size decrease corresponds to a 10,000 times higher intensity at the sample with the same laser power. The collinear spectroscopies

presented in this thesis thus have the potential to acquire the same results as the most advanced nonlinear spectroscopies with less expensive and smaller laser sources than are typically required.

1.2 Gradient Terms become Relevant

In Chapter V we explore the effects of a steep optical gradient on quasiparticle excitons in semiconductor quantum wells. Optical tweezers are commonly used in biology labs to move and control cells, and optical traps are used in atomic physics to control atoms. Here we find that excitons are also susceptible to similar gradient terms. Since electron-hole pairs have very small dipole moments (≈ 5 Åe), the gradient effects are reliant on many-body effects that enhance the gradient terms.

This means that we have found a regime at which the homogeneous semiconductor Bloch equations, the standard equations for describing the optical response of a semiconductor, are not sufficient. This project has been a collaboration with theorists Prof. Mackillo Kira and his student Markus Borsch. In the process of developing a theory to describe the optical gradient effects that we measure, they have developed solutions and methods for treating the inhomogeneous semiconductor Bloch equation. In this thesis we explore an inhomogeneous excitation by a diffraction-limited spot, but the theories can be extended to inhomogeneous semiconductor systems.

1.3 Spatially Inhomogeneous Samples are Accessible

Not all samples are homogeneous over a large excitation spot. For very short-range disorder, much smaller than the wavelength of the light used to measure it, the inhomogeneity must be measured with near-field techniques or advanced optical techniques like multidimensional coherent spectroscopy (MDCS). For longer-range inhomogeneity, on the order of microns, it is sufficient to focus tightly. In this thesis I discuss our exploration of samples having all kinds of inhomogeneity with varying

length scales. Having a diffraction-limited spot has enabled a number of our experiments because it has allowed us to probe local sample regions. On top of using a small spot, we have developed a collinear MDCS technique that allows us to measure the inhomogeneity at smaller length scales than the diffraction limit.

In Chapter VI we distinguish individual interfacial quantum dots (QDs), point-like confined excitations of a semiconductor, by probing a sample region that is sufficiently small to spatially isolate the QDs. We probe these QDs with MDCS to measure interactions between QDs, and we demonstrate how we can induce those interactions. With a larger spot, QDs in this system appear as a continuum of states. Our ability to isolate individual QDs wound up being essential for understanding how interactions between QD states were induced and for showing how it could be useful. Since the QDs are localized to regions that are smaller than 100 nm, we also discuss how we can study physics at the sub-diffraction-limited length scale by isolating QDs and measuring the timescales of their interactions.

In Chapters VII and VIII we measure monolayer transition metal dichalcogenide (TMDC) samples. The samples are made by exfoliation, which is the technique known to make the highest quality monolayers. Exfoliated samples are typically only a few microns in size, and so most of these samples cannot be measured with non-collinear techniques. Our ability to measure these samples with MDCS has enabled us to determine their homogeneous linewidths and use the homogeneous linewidth as a metric to improve the samples. Since some of these samples have only been measured with linear techniques, such as photoluminescence, we have been able to assuage some misconceptions about the disorder in the encapsulated monolayer samples. We have also used MDCS to measure novel monolayer TMDC structures with exceptional coherence properties that have been created in Prof. Hui Deng's group by Jason Horng.

In Chapter IX we use MDCS to measure confined exciton-polaritons and their

various couplings. These samples are laser devices created in Hui Deng's lab by Bo Zhang. The devices are 7.5 μm squares with polariton excitations that are confined to much smaller regions than the device size. We measure coupling between different polariton modes, and we use the spatial selectivity of our tightly focused beams to pump and probe at spatially different locations. We also measure relaxation from the exciton states into the polariton states. When these devices are used for lasing, the exciton states are pump and the polariton states lase. We are therefore able to directly measure those relaxation pathways.

We finally discuss how the coherence length of quasiparticle excitations affects a sample response. By measuring a sample with a coherence length that is greater than the excitation spot size, we show that it is possible to induce nonlocal nonlinear wave mixing. In Chapter X we discuss a theoretical treatment of this phenomenon, and we present a first experimental demonstration of nonlocal four-wave mixing.

CHAPTER II

Heterodyne Detected Resonant Nonlinear Spectroscopies

2.1 Introduction

Collinear resonant spectroscopic techniques all have a major technical challenge to overcome: distinguishing the signal from everything else. Here I discuss the solution to this problem for nonlinear spectroscopy. The solution, heterodyne detection of frequency-shifted beams, is a coherent technique, but it does not require any of the exciting light sources to necessarily be lasers. We apply this heterodyne technique to pump-probe spectroscopy with a high-order of nonlinearity, and we apply it to a set of phase-resolved third-order nonlinear spectroscopies called multidimensional coherent spectroscopy (MDCS). Throughout this thesis we apply these techniques to semiconductor samples, but they can also be applied to measuring atoms [13], molecules [14], and biological samples [15].

In this chapter I begin with a fundamental introduction to linear and nonlinear spectroscopies. I then introduce transient absorption spectroscopy. With this spectroscopy as an example, I can concretely describe how we use heterodyne detection to isolate the transient absorption signal from the two pulsed beams used to generate that signal.

2.2 Linear Absorption Spectroscopy

A light source will induce a polarization in any polarizable material. Since these polarizations result from oscillating charge, we can describe these excitations on a most fundamental level with a Lorentz oscillator model. The model is the solution to a driven oscillator with damping. Since we are typically measuring many charges, we apply the driven oscillator model to a polarization vector, which is defined as the density of dipole moments. The polarization of a material with density N of oscillators each with charge q is $\mathbf{P} = Nq\mathbf{x}$. The damping rate γ of the oscillator results from internal scattering and photon emission. The mass m_e is the effective mass of the oscillating charge, which is typically an electron bound to a heavy nucleus. The natural frequency of the oscillator ω_0 corresponds to the binding energy of the oscillator. The driving force is the product of the particle charge and the applied electric field, which is time dependent. The particle is typically an electron with fundamental charge $q = -|e|$ and the field is a sinusoid describing the excitation light $\mathbf{E}(t) = \mathbf{E_0}e^{-i\omega t}$. Since the excitation field is sinusoidally varying, we assume a form of the polarization that is also sinusoidally varying: $\mathbf{P}(\omega, t) = \widetilde{\mathbf{P}}(\omega)e^{-i\omega t}$. $\widetilde{P}(\omega)$ is thus independent of t, but it has a phase that depends on ω. We find the driven polarization is

$$\begin{aligned}\widetilde{\mathbf{P}}(\omega) &= \frac{Nq^2/m_e}{\omega_0^2 - \omega^2 - i\gamma\omega}\mathbf{E_0} \\ &= \frac{Nq^2}{m_e}\left(\frac{\omega_0^2 - \omega^2}{\left(\omega_0^2 - \omega^2\right)^2 + (\gamma\omega)^2} + i\frac{\gamma\omega}{\left(\omega_0^2 - \omega^2\right)^2 + (\gamma\omega)^2}\right)\mathbf{E_0},\end{aligned} \tag{2.1}$$

where ω is the frequency of the driving field. We write the response as a function of frequency and note that it has real and imaginary components that correspond to refraction and absorption, respectively. The polarization induced is generally written: $\mathbf{P}(\omega) = \epsilon_0\chi(\omega)\mathbf{E}(\omega) = \epsilon_0[\chi'(\omega) + i\chi''(\omega)]\mathbf{E}(\omega)$, where χ' and χ'' represent the real and imaginary parts of the dielectric susceptibility. It is helpful to introduce the

measurable complex refractive index to relate the effect of a material's susceptibility on a wave propagating according to the wave equation. The wavevector k of a wave propagating through a material with complex refractive index $\widetilde{n}$ is $k = \frac{\omega}{c}\widetilde{n}$. The complex refractive index is decomposed into real and imaginary parts, called the refractive index n and the extinction coefficient κ, respectively. These quantities are related to susceptibility as follows: $\widetilde{n} = n + i\kappa = \pm\sqrt{1 + \chi' + i\chi''}$. Now we can write down a field propagating through a material:

$$\mathbf{E}(z) = \mathbf{E_0} e^{i\frac{\omega}{c}(n+i\kappa)z} = \mathbf{E_0} e^{-\frac{\omega}{c}\kappa z} e^{i\frac{\omega}{c}nz}, \tag{2.2}$$

where z is the direction of propagation. We now relate this equation to Beer's law for linear absorption. Beer's law is derived by considering the decrease in intensity of light across an incremental slice of thickness dz due to absorption defined by the absorption coefficient α: $dI = -\alpha I(z)dz$. By integrating this we obtain Beer's law: $I(z) = I(0)e^{-\alpha z}$. We immediately see by relating Beer's law to the absolute square of Eq. 2.2 that $\alpha = \frac{2\omega}{c}\kappa$.

For low intensity light the dielectric susceptibility is constant, so the induced polarization scales linearly with the field. With a tunable monochromatic light source, or by spectrally resolving the transmission of white light through a sample, one can thus measure the absorption coefficient. For a dilute medium, $\chi'' \ll 1$. We find in this limit that

$$\kappa = \frac{1}{\sqrt{2}}\sqrt{-1 - \chi' + \sqrt{(1+\chi')^2 + \chi''^2}} \approx \chi''. \tag{2.3}$$

In solid state systems this approximation is not always valid, and it is necessary to calculate local field corrections [16].

2.3 Nonlinear Optical Response

High intensity excitation light can drive a material response into a nonlinear regime. The nonlinear response is typically written as a perturbative expansion:

$$\mathbf{P}(\omega) = \epsilon_0(\chi^{(1)}\mathbf{E}(\omega) + \chi^{(2)}\mathbf{E}(\omega)^2 + \chi^{(3)}\mathbf{E}(\omega)^3 + ...), \tag{2.4}$$

where $\chi^{(n)}$ are $(n+1)$-th-rank tensors that represent the polarization-dependence of the parametric interaction and the symmetries of the material. This treatment of nonlinear signals is a powerful way to realize and describe nonlinear wave mixing. For instance, we can realize that even-order nonlinearities (for instance $\chi^{(2)}$) vanish in inversion symmetric materials. Since the induced polarization in an inversion symmetric material must change sign if the electric field changes sign, even order susceptibilities must be zero. $\chi^{(3)}$ is thus the lowest-order nonlinear term with no symmetry restrictions.

It is not essential to treat a sample response perturbatively. Full solutions to the equations that govern light-matter interaction can be used. Full solutions to the optical Bloch equations, for instance, introduce Rabi flopping in atoms excited with strong fields. The semiconductor Bloch equations accurately describe the response of a semiconductor up to complete saturation of the exciton resonance [17]. In order to measure the response of a material that includes high-order responses, we use pump-probe spectroscopy. Ultrafast pulses have very high peak intensities that more easily enable access to highly nonlinear regimes. The nonlinear signal in a pump-probe geometry, described below, is radiated in the direction of the probe pulse. While phase matching conditions cause the emission direction to depend on the nonlinear order of the signal, all orders are measured in the direction of the probe.

2.4 Wavevector-selected Transient Absorption

In a resonant pump-probe experiment two pulses are used to generate a nonlinear signal. The first pulse to excite the sample is called the pump. A resonant pump pulse will induce a polarization in the sample, and in semiconductor nanostructures it will typically create a population of bound electron-hole pairs called excitons. The pump pulse can be very nonlinear. The second pulse is designed to probe the response of the sample some finite delay after the sample is excited by the pump. Typically the probe will be in a linear regime where the polarization it induces scales linearly with the probe field strength. To determine the nonlinear signal, one will measure the difference of the probe response with and without the pump excitation. This measurement is called differential pump-probe or differential absorption. When the signal is measured as a function of the delay between the pulses, it is called transient absorption.

In order to measure the nonlinear signal with resonant beams, the pump and probe beams typically have different wavevectors so they may be distinguished. To spectrally resolve a transient absorption signal, a monochromator and photodetector are used to measure the light in the probe direction. Assuming no light from the pump is scattered into the detection wavevector, the differential intensity measured on the detector for frequency ω set by the monochromator is

$$\Delta I(\omega, T) = |E_{\text{probe}}(\omega) + E_{\text{NL}}(\omega, T)|^2 - |E_{\text{probe}}(\omega)|^2, \tag{2.5}$$

where E_{probe} is the field of the transmitted probe beam that includes effects of the induced linear polarization, E_{NL} is the nonlinear response, and T is the temporal delay between the pump and probe beams. In the limit where $E_{\text{NL}} \ll E_{\text{probe}}$, $\Delta I(T) \approx 2|E^*_{\text{probe}} E_{\text{NL}}(T)|$. This corresponds to heterodyne detection of the nonlinear signal with the probe.

This subtraction is further complicated when weak pump scattering is also considered. It is essential to also measure the scattered pump signal and subtract that from the measured nonlinear signal. Subtracting linear spectra to determine a weak nonlinear spectrum is fairly noisy, so we prefer to do all the subtraction dynamically. One method for distinguishing the signal from the pump and probe beams is to independently amplitude modulate the pump and probe beams. Amplitude modulation is typically performed by a mechanical chopper at a few kHz, but there are ways of amplitude modulating a beam at much higher frequencies too. If pump and probe beams are amplitude modulated with frequencies ω_{pump} and ω_{probe}, the differential signal resulting from the nonlinear interaction between the beams will be amplitude modulated at $\omega_{\text{pump}} \pm \omega_{\text{probe}}$. For quantitative measurements, and in order to capture all orders of nonlinearity, it is best to amplitude modulate with a square wave. One can do this by chopping the beam at a focus. We see then that the sum and difference frequencies are proportional to the differential signal. By lock-in detecting at the nonlinear signal tagging frequency it is possible to extract the nonlinear signal from the other beams impinging the detector. Since the lock-in amplifier filters just the fundamental frequency of the square signal, the coefficient of proportionality is determined by truncation of the higher harmonics. For a detailed treatment of quantified transient absorption see the Appendices of [18].

The spectrally-resolved version of transient absorption, where the beam is spectrally filtered before the detector by a monochromator, is commonly used to measure decay times of various spectral features. A very careful implementation of spectrally-resolved transient absorption has been used to quantify the excitation density dependence of many-body effects in GaAs quantum wells [8] and to discover quantum droplets and determine their lifetime [19].

2.5 Heterodyne-detected Transient Absorption

If the scattered pump light becomes comparable to the probe beam, this differential method breaks down. This case is clearly realized for collinear pump and probe beams where the differential signal is:

$$\Delta I(\omega,\, T) = |E_{\text{probe}}(\omega) + E_{\text{pump}}(\omega) + E_{\text{NL}}(\omega,\, T)|^2 - |E_{\text{probe}}(\omega)|^2 - |E_{\text{pump}}(\omega)|^2. \quad (2.6)$$

With a significant amount of pump light in the beam, the nonlinear signal is heterodyned with both the pump and the probe, and the T delay affects the interference of the pump and nonlinear signal that is emitted after the probe. We thus need some way to isolate the nonlinear signal field such that we do not measure terms corresponding to the interference of that field with any of the pulses acting on the sample.

We accomplish the measurement of E_{NL} by heterodyne detecting the field with an external local oscillator (LO) pulse that we route around the sample, shown in Fig. 2.1. This technique was first introduced by K. L. Hall *et al.* for performing resonant pump-probe spectroscopy in waveguides [20]. Specifically, instead of amplitude modulating the probe beam, we shift the frequencies of the probe and the LO so that the interference between them is amplitude modulated at a frequency that is unique to that interference term. The beams are an amplitude modulated pump (by square wave with frequency ω_{pump}), a frequency shifted probe (by ω_{probe}), and a frequency shifted LO (by ω_{LO}). The fields of the various beams are:

$$\begin{aligned} E_{\text{pump}}(t) &= E_{\text{pump, no mod.}}\, \text{sgn}(\sin(\omega_{\text{pump}} t)) e^{i\omega_{laser} t} \\ E_{\text{probe}}(t) &= |E_{\text{probe}}| e^{-i(\omega_{\text{laser}} + \omega_{\text{probe}})t} \\ E_{\text{LO}}(t) &= |E_{\text{LO}}| e^{-i(\omega_{\text{laser}} + \omega_{\text{LO}})t}, \end{aligned} \quad (2.7)$$

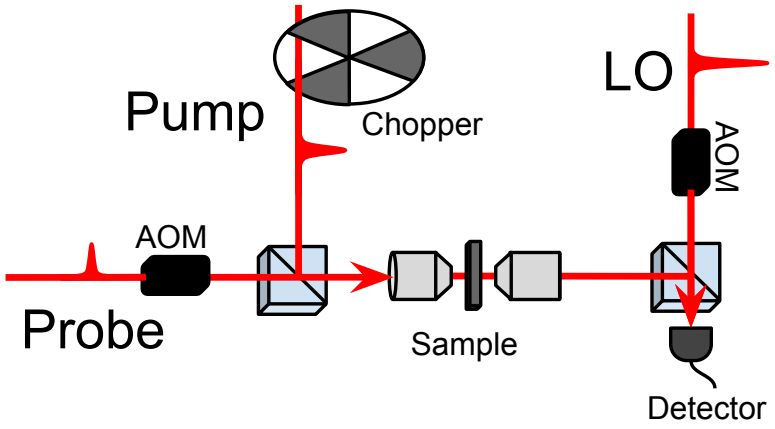


Figure 2.1: Nonlinear signal is generated by amplitude modulated pump beam and frequency shifted probe beam. Amplitude modulation can be performed with a mechanical chopper and frequency shifting is performed with acousto-optic modulators (AOMs). A local oscillator (LO) beam, which is frequency shifted by a different frequency than the probe beam, interferes with the signal on a detector. Since the modulation on the detector corresponding to the interference between the pump-probe signal of interest and the LO is unique, we can isolate the signal with a lock-in detector tuned to that modulation frequency.

where ω_{laser} corresponds to a narrow frequency component of the light source used to generate the probe and LO. We use an ultrashort pulsed laser with relatively high repetition frequency, but we could also use incoherent light or a continuous-wave (CW) laser. The restrictions on the light source that can be used for frequency-shifted heterodyne detection will become clear below. Now we see that if we interfere the probe field and the LO field on a detector, without considering the effect of the pump, we will measure $I = |E_{\text{probe}}|^2 + |E_{\text{LO}}|^2 + E^*_{\text{LO}} E_{\text{probe}} e^{-i(\omega_{\text{probe}} - \omega_{\text{LO}})t} + \text{CC}$. The linear interference between the probe and LO is modulated at the difference in their frequency shifts, which we typically set to be on the order of 1 MHz. The amplitude modulated pump will introduce the nonlinear signal in the frequency domain as sidebands on the probe field. Ignoring the signals at DC, we will find that there are three sources of modulated signals, called beat notes, on the detector. 1) Odd harmonics of the square pump modulation will be measured at low frequencies. 2) There are a few modulated linear interference terms on the detector. Of course the linear interference

of the probe and LO will lead to a beat note at the difference in their frequency shifts. If the pump is created by the same source as the probe and LO, there will be linear interference terms between the latter two beams and all harmonics of the pump beam. To minimize this effect we will typically carefully ensure the frequency shifts of both the probe and LO are significant relative to the pump so that those interference terms are outside the detector bandwidth. 3) The last frequencies that will appear on the detector are sidebands on the probe-LO beat note at the odd harmonics of the pump modulation frequency. These sidebands correspond to the radiated non-linear signal resulting from the pump and probe interfering with the LO. We will typically measure at the fundamental sideband, which corresponds to the beat notes at $\omega_{\text{probe}} - \omega_{\text{LO}} \pm \omega_{\text{pump}}$. These beat notes can be uniquely attributed to a radiated four-wave mixing (FWM) signal that is interfered with the LO on the detector.

For frequency shifting to work it is essential for the frequency shifted probe to be distinguishable from the LO. Pulsed lasers with relatively high repetition rates have few restrictions on how they can be shifted. Though the bandwidth of the pulsed laser is large, the spectrum is made up of a narrow near-delta functions spaced by the repetition frequency of the laser. In a stabilized laser, these are called comb teeth, and they have a linewidth corresponding to the coherence length of the laser. Though the lasers we use are not stabilized, the probe and LO are generated from the same laser, and thus the laser noise is common between them. We can thus relate that the heterodyne signal will have the underlying frequency comb spectrum with a comb-tooth spacing of the repetition frequency, shown in the left panel of Fig. 2.2. For frequency shifts that are very near a multiple of the repetition rate, the shifted spectrum becomes nearly indistinguishable from the unshifted spectrum. The heterodyne technique is thus restricted for low repetition rate lasers in which finding an acceptable shift frequency is non-trivial. The random phase across the spectrum of an incoherent light source allows arbitrarily frequency shifted spectra

to be distinguishable from their unshifted counterparts, shown in the right panel of Fig. 2.2. It is necessary for the incoherent source to be spatially coherent, since the fields must coherently interfere. The resolution will be limited by the detection technique, which is typically determined by the scan duration of a Fourier transform spectrometer. In summary, the frequency shifting technique is restricted when there is high spectral coherence and low temporal coherence.

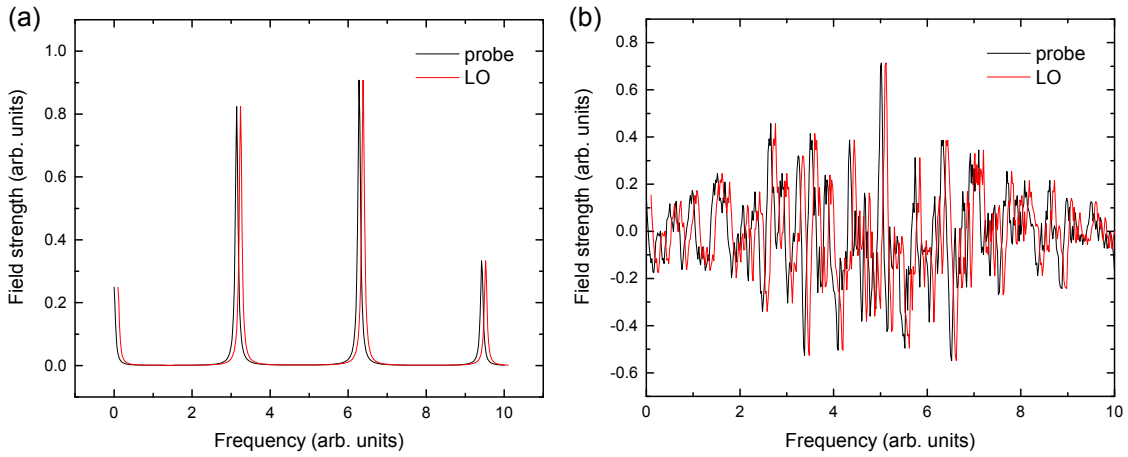


Figure 2.2: Frequency shifted spectra that will both have resolved beat notes between the labeled probe and LO beams if the two are interfered on a detector. (a) A pulsed laser will have a frequency comb spectrum, where correlated noise is subtracted. If the frequency shift is greater than linewidth of the detected repetition rate and not equal to a harmonic of the repetition rate, the beat note between the probe and LO will be resolved and non-zero. (b) A light source with a random spectral phase, e.g. a thermal source, will also have a resolved beat note for any frequency shift the detection technique is capable of resolving. There is no fundamental limit imposed by the incoherent source so long as it is coherent to first order.

We now have two relative pulse delays that we can scan. In Fig. 2.3a we plot the nonlinear absorption signal of a resonance in a gallium arsenide quantum well as a function of the LO delay with respect to the probe pulse. Here we measure the free-polarization decay, which is the time-domain representation of the spectrum. The time constant of the free-polarization decay is the inverse of the linewidth. We can integrate over this decay for the equivalent of what one would measure with differential absorption. In Fig. 2.3b we plot the integrated differential absorption

signal as a function of the delay between the pump and the probe pulses. This is the equivalent of a transient absorption experiment that measures the relaxation of the population created by the pump pulse.

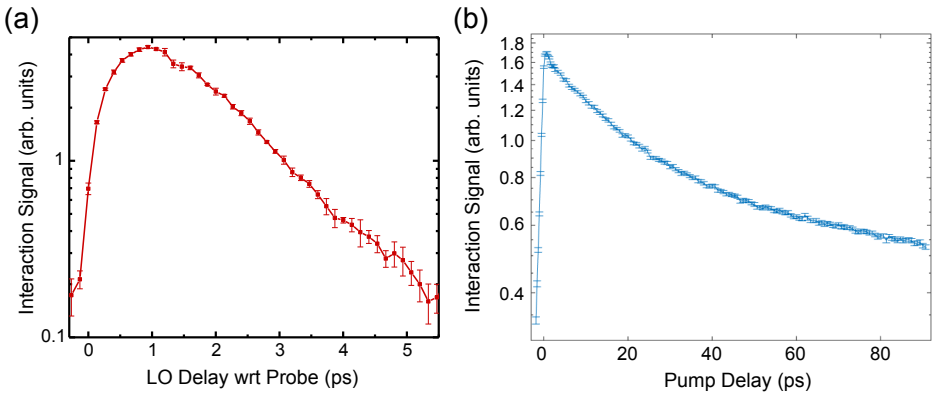


Figure 2.3: Nonlinear transient absorption measurement of a resonance in gallium arsenide quantum wells. (a) We measure the signal as a function of the delay between the LO and the probe. This signal corresponds to the free polarization decay, which has a timescale that is the inverse of the resonance linewidth. (b) We measure the integrated free polarization decay as a function of the delay between the pump and probe pulses. This is a transient absorption measurement and reveals the timescale of the population decay.

2.6 Fourier-transform Spectroscopy

In the previous section we mention an inverse relationship between our measurement of the time-domain decay of a nonlinear signal and the spectral linewidth of that signal. These two domains are of course related by a Fourier transform, but it is necessary to measure the signal phase to go between these domains. Fourier transform spectrometers are commonly used to measure infrared light where the stability requirements are not as stringent as in the optical part of the spectrum. Recently optical Fourier-transform spectrometers have become more common with the development of various phase stabilization and phase retrieval methods.

To introduce the advantages of Fourier-transform spectrometers, I will compare them to conventional dispersion spectrometers. Conventional spectrometers disperse

light on a grating so that a detector will measure a single wavelength of the signal at a time. The wavelength that hits the detector can be scanned to measure the spectrum in a device called a monochromator, which is fairly slow. Alternatively an array detector can be used to simultaneously measure many spectral components at once. Fourier-transform spectrometers measure the time-domain interference of two coherent beams. These beams must be in some way a copy of each other generated in an interferometer, though one copy is commonly passed through a dispersive or absorptive medium so that its spectrum differs from the other pulse. The interference of the two beams is measured as a function of the delay between the interferometer arms, which can make them slower than a spectrometer with a detector array. Comparing the two spectrometer types, we see that the Fourier-transform spectrometers are typically less expensive since they do not require a grating or a detector array. The other major advantage that was pointed out by P. B. Fellgett in his thesis is that multiplexed measurements have a higher signal-to-noise ratio when detector noise is considerable [21]. When detector noise is higher than the shot noise of a measurement, using more detectors to measure the same total signal increases the noise proportionally to the number of detectors. Though this is typically not an issue in optical measurements where we are shot-noise limited, infrared detectors are notoriously noisy. The last issue with detector arrays is that each detector must be processed independently. Lock-in detecting an array of pixels requires either an array of lock-in amplifiers or careful parallel processing of the detectors' outputs. Since we desire to use lock-in detection to distinguish the nonlinear signal in our experiment, the choice to spectrally resolve the signal by Fourier-transform spectroscopy is clear.

In order to implement the optical Fourier-transform spectrometer we must measure both the amplitude and phase of the signal at each delay between the reference and signal pulses (which are both generated from the same source). Measuring the phase of visible light is not trivial because the frequencies are so high. We typically

work with light that has a wavelength of 800 nm, which means that typical relative path length fluctuations on the order of 1 μm lead to phase fluctuations that are more than two orders of magnitude greater than acceptable 1% fluctuations. Active stabilization with loop filters and piezoelectric controlled mirrors has been used to demonstrate $\lambda/100$ stability in three interferometers [22, 23]. A passively stable interferometer for heterodyne detection has also been demonstrated using a single diffractive optic to generate the beam copies and either common or transmissive optics otherwise [24]. There is, however, precedent for another set of solutions that instead only rely on measuring the fluctuations.

The typical use of a reference laser to measure a signal with an unstable interferometer is in infrared spectrometers, where the light has a wavelength between 2000 and 20000 nm. Since the instability of the interferometers is less than the wavelength of infrared light, it is possible to know the delay within one wavelength. To determine the phase beyond that, it is standard to use a He-Ne laser having a wavelength of 633 nm as a reference for spectrally resolving the light. In this case, one just needs to scan the relative delay between the interferometer arms and use the zero-crossings of the reference laser interference to identify points spaced by half the reference wavelength. Since the interferometer is referenced with a relatively short wavelength laser, the sampling spacing is known with high interferometric precision. This solution does not scale to the visible spectral region for a couple of reasons. For reference lasers that have a similar wavelength or a much longer wavelength than the signal, it is not possible to sample the signal adequately with zero-crossings. Additionally, if the natural fluctuations of the interferometer are greater than λ, zero-crossings in the backwards direction unintentionally trigger a measurement by this technique. One method for solving the problem of zero-crossing ambiguity is use in-quadrature optical detection to measure the interferometer output. This technique requires two detectors and additional optics, but it can be used to distinguish 0 and π phase zero

crossing.

For visible light in a frequency shifted interferometer, a different referenced detection is used to determine the relative phase between two interferometer arms. Frequency shifting one arm with respect to the other dynamically cycles the phase of the interference between the two arms on a detector. Basically this frequency shift permits the measurement of the signal with many different phases on a timescale that is much faster that the path length fluctuations of the interferometer. By comparing the modulated signal to a modulated reference laser in the same interferometer, shown in Fig. 2.4, it is possible to retrieve the signal phase and amplitude at every sampling point. The process of demodulating the signal is called synchronous sampling, and it is a common technique used in radio-frequency electronics. This technique is primarily used in optical interferometers for Fourier-transform spectroscopies [13, 25, 26].

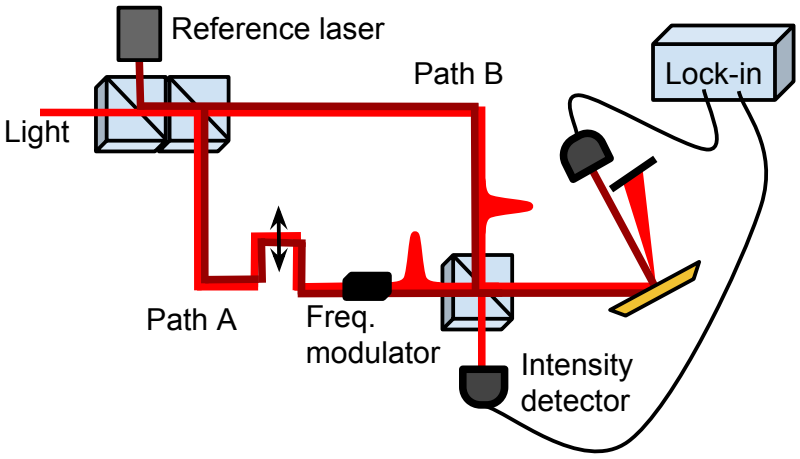


Figure 2.4: Schematic for an frequency shifted interferometer for spectroscopy. A reference laser samples the same path length fluctuations as the signal light. The signal amplitude and phase are realized by synchronous detection using a lock-in amplifier.

In the interferometer in Fig. 2.4, a reference laser copropagates with the light source of interest (the signal light) to sample all the same fluctuations as the light. The reference laser is separated from the signal by either spectral filter or polarization,

depending on how the light is distinguished. The interfered reference beams and signal beams are measured on separate detectors. The interference signal is measured by demodulating the signal with the reference, typically using a lock-in amplifier. If we write down the phase of each interference we see some wavelength dependence:

$$\begin{aligned}\phi_{\mathrm{ref}} &= \phi_{\mathrm{ref},\,0} + \frac{2\pi}{\lambda_{\mathrm{ref}}}(x_{\mathrm{t}} + \delta_{\mathrm{fluc}}) + f_{\mathrm{mod}}t, \\ \phi_{\mathrm{sig}} &= \phi_{\mathrm{sig},\,0} + \frac{2\pi}{\lambda_{\mathrm{sig}}}(x_{\mathrm{t}} + \delta_{\mathrm{fluc}}) + f_{\mathrm{mod}}t,\end{aligned} \tag{2.8}$$

where x_{t} is the intended relative path length (corresponding to a time delay), δ_{fluc} is the path length difference due to flucutuations of the interferometer, and f_{mod} in the relative frequency modulation of the two arms. The quantity measured by synchronous sampling of the interferometer corresponds to $\phi_{\mathrm{sig}} - \phi_{\mathrm{ref}}$. The resulting phase still fluctuates with the interferometer, but that fluctuation is suppressed by $|\lambda_{\mathrm{sig}}/\lambda_{\mathrm{ref}} - 1|$. A reference having the same wavelength as the signal will perfectly measure phase fluctuations of the interferometer and cancel those from the signal.

The phase also evolves with the time delay at the frequency difference between the signal and reference. When we Fourier-transform the signal with respect to that delay, we measure the signal frequencies with respect to the reference frequency. In Fig. 2.5 we plot a measured time-domain signal and corresponding frequency-domain signal of a GaAs quantum well resonance. Here the nonlinear signal is generated in one arm of the interferometer and interfered with a LO pulse in the other arm. Before Fourier transforming the time-domain signal we zero-pad the data by a factor of two, which doubles the number of points in the spectrum. Generally the interferogram length should always be at least doubled by zero-padding to accurately reflect the resolution of the technique [27]. Additional padding will smooth the data and not actually affect the instrument response, but it does not increase the resolution. We then Fourier-transform the signal using a fast Fourier transform.

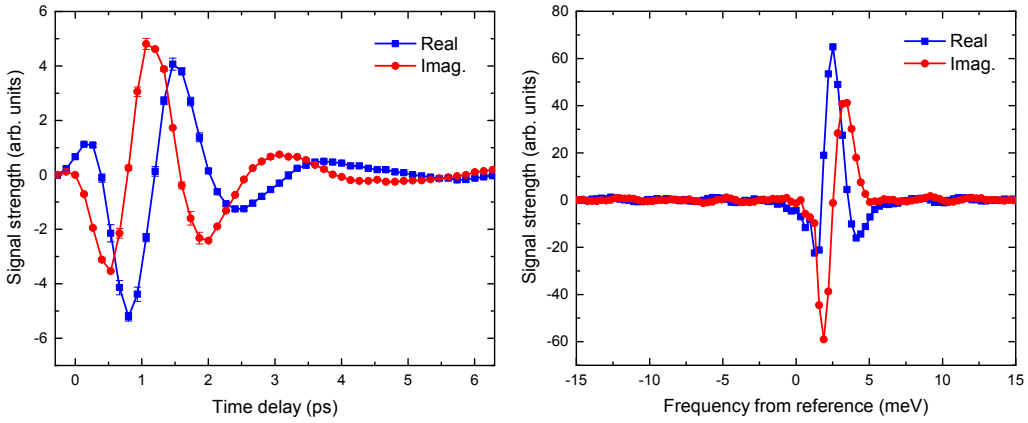


Figure 2.5: Measured time-domain and calculated frequency-domain nonlinear signal for the same GaAs resonance plotted in Fig. 2.3. The signal is measured with respect to a reference at a wavelength of roughly 804 nm, or 1542 meV.

Another advantage of time-domain spectroscopy is its tremendous versatility. By careful selection of the sampling rate and scan duration it is possible to tailor the measurement to the sample of interest. Though the measurement in Fig. 2.5 is sufficient to realize the spectrum of a GaAs quantum well resonance, it is an example of a mediocre selection of scan parameters. The time-domain signal is truncated by the insufficient scan duration. The spectrum also has an excess of bandwidth due to the small step size of the measurement. In measurements throughout this thesis, parameters are instead selected to optimally sample the spectrum. The step size is chosen so that the bandwidth of the measurement is just sufficient to capture all of the features. The scan duration extends until the signal is fully decayed to within the measurement noise so that the measurement resolution is optimal.

For measurements where the reference frequency is relatively far detuned from the signal we undersample the signal to improve the acquisition time. Frequencies that oscillate faster than the sampling rate are down converted to a frequency within the measurement bandwidth. For a phase-resolved measurement the bandwidth is identical to the sampling rate, which is $f_s = 1/t_{\text{step}}$ in an ultrafast measurement. Therefore for a sampling rate corresponding to 30 meV, the range [-15, 15) meV is

identical to [15, 45) meV. We often use undersampling throughout this thesis to reduce the sampling rate and thus decrease scan durations. It is important to note that by quadrature detecting the signal using synchronous sampling we significantly simplify undersampling. Measurements that measure only one quadrature of the signal cannot distingush the evolution of positive and negative frequencies.

2.7 Multidimensional Coherent Spectroscopy

Semiconductor nanostructures, like the GaAs QWs introduced above, have significant inhomogeneity introduced by spatial fluctuations of the nanostructure size in the confinement direction. We will describe the sources of inhomogeneity in Chapter III, but here it is sufficient to explain that inhomogeneity broadens resonances in optical spectra. In linear spectroscopy it is difficult to infer the intrinsic linewidth of a resonance when the measured lineshape is a sum of measurements of many resonances with different center energies. Inhomogeneity limits the capabilities of optical spectroscopy of other systems as well: atomic lines are inhomogeneously broadened by Doppler broadening, and molecular spectra are typically cluttered by the response of other molecules that contribute additional features to the spectra. Multidimensional coherent spectroscopy (MDCS) offers a solution to measuring the intrinsic structure of semiconductor, atomic, and molecular systems in the presence of inhomogeneity.

By also spectrally resolving the pump absorption in the above transient absorption experiments, we are able to measure correlations between resonances. Basically the nonlinear signal of a resonance is measured as a function of the absorbed photon energy and the emitted photon energy. A single oscillator will absorb and emit light at the same energy. An ensemble of uncoupled oscillators will all absorb and emit at their respective energies, but there will be no signal resulting from absorption at one energy and emission from another. This is evident in an MDCS measurement of GaAs QWs plotted in Fig. 2.6. In this spectrum, measured by Alan Bristow et

al. using a wave-vector selected MDCS technique, measured resonances are aligned along the diagonal of the spectrum where |absorption photon energy| = |emission photon energy| [1]. Cross-diagonal slices correspond to the homogeneous linewidth of the individual oscillators, and the diagonal slice correspond to the inhomogeneous distribution of those oscillators. MDCS also reveals coupling between two resonances. In Fig. 2.6, the two resonances correspond to two non-degenerate valence bands in the GaAs QWs. The nonlinear signal reveals coupling between these two resonances. The small elongation along the diagonal direction of the off-diagonal features reveals that the inhomogeneity of the two resonances is correlated; that is, high energy oscillators in the X_{hh} resonance couple to high energy oscillators in the X_{lh} resonance.

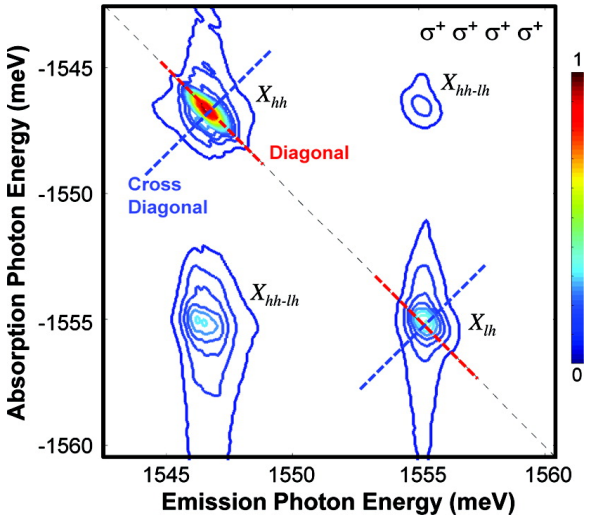


Figure 2.6: Multidimensional coherent spectroscopy measurement of two resonances in GaAs QWs: X_{hh} and X_{lh}. The spectroscopy correlates absorption and emission photon energies revealing homogeneous linewidths of the two resonances despite sample inhomogeneity. The spectrum also indicates there is coupling between the two resonances with correlated inhomogeneity. The absorption energies are negative because this is a rephasing-type measurement. The measurement was made with co-circularly polarized excitation and detection beams. Figure from [1].

In the collinear geometry, we resolve the absorption axis using two pump pulses instead of one. We split the pump pulse into two in an interferometer. To generate, delay, and phase-resolve the pump-pulse pair we use the same scheme as presented in Fig. 2.4. Rather than going to the intensity detector, the pump pair is combined with

the probe beam and directed to the sample of interest. Now to distinguish the signal, we uniquely frequency shift both pump beams (labeled A and B), the probe beam (labeled C), and the LO by ω_A, ω_B, ω_C, and ω_{LO}, respectively. The interference of beams A and B at the sample will be modulated at the difference in their frequencies: $\omega_B - \omega_A$. This modulation is an intensity modulation that resembles the amplitude modulation of the pump discussed earlier to measure transient absorption. We can thus identify the interference of the nonlinear MDCS signal and LO at the frequencies $\omega_C - \omega_{\mathrm{LO}} \pm (\omega_B - \omega_A)$. These two frequencies have different meanings that we will address momentarily. We can now measure the nonlinear signal with a frequency-resolved absorption axis. Since we scan the delay between pulses A and B to generate the spectra, the T delay where population relaxation occurs is the temporal delay between pulses B and C.

Since the MDCS techniques presented here use ultrafast pulses to generate high resolution spectra, the spectroscopy is at an exciting intersection of having high temporal and energetic resolution. We can exploit this intersection to measure population relaxation dynamics [28] and spectral diffusion [29], which can both reveal valuable information about a system's electronic dynamics.

To simplify the interpretation of MDCS, the interaction of pulses with the sample is typically treated perturbatively. We therefore set the power of all excitation beams (A, B, and C) to be equal and relatively weak. Most commonly used MDCS techniques measure third-order nonlinearity. We are careful to measure the nonlinear regime we are in by determining the scaling of the nonlinear signal with pulse power. For excitation with three pulses, we expect the signal to have a cubic scaling with the excitation fields. Once we verify that we are in the third-order nonlinear regime, signals excited by exactly three field interactions will dominate signals excited by five or seven field interactions. If we measure a signal resulting from the nonlinear interaction of all three excitation beams (A, B, and C) we can identify how the beams

interacted with the sample to generate the nonlinear signal.

The MDCS pulse sequence used for measuring homogeneous linewidths of inhomogeneously broadened system is just one of many useful pulse sequences. This one is called the rephasing pulse sequence because it corresponds to the coherent phase evolution between pulses A and B being opposite the phase evolution of the signal. In Fig. 2.7 we denote that pulses A and LO are conjugated pulses. Pulse A^* polarizes the sample, and this field interaction is typically called "creating a coherence". This coherence evolves for time τ. Pulse B then freezes the coherence to excite the system into a population state. Earlier we related this to the interference of A and B at the sample, which is an equivalent way to think of the system after this field interaction. The population evolves for time T before interacting with pulse C. Pulse C again creates a coherence in the system, but since C is not conjugated the coherence evolves opposite the coherence created by A. The evolution of the signal after pulse C is measured with a LO pulse, which is delayed t with respect to pulse C. At this point one might ask what happened to the signals due to all the other field interactions with the sample. Those all also occur in the sample, but their responses are tagged with different modulation frequencies after being interfered with the LO on the detector. When we measure the signal at the frequency $-\omega_A + \omega_B + \omega_C - \omega_{\mathrm{LO}}$, we are selecting the signal corresponding to rephasing.

We can understand how the rephasing pulse sequence gets its name by considering the effect of the pulse sequence on an inhomogeneously broadened system in the time domain. Pulse A will excite the inhomogeneous distribution of oscillators, and they will all begin evolving at different rates. The coherent response will dephase according to the linewidth of the inhomogeneous distribution. After dephasing for time τ, Pulse B and C will freeze and then reverse the evolution of the oscillator distribution. Oscillators evolving at high (low) rates will still evolve at a high (low) rate in the backwards direction. After evolving for time $t = \tau$, the oscillators will

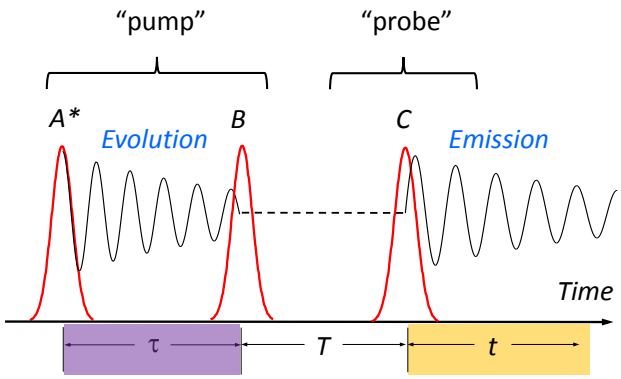


Figure 2.7: MDCS rephasing pulse sequence. The pump is split into two pulses that create an exciton population. The pump absorption is spectrally resolved by scanning the time delay between A* and B and Fourier transforming the respons with respect to that delay. The probe pulse, C, stimulates emission of the MDCS signal that is resolved by heterodyne detection with a fourth LO pulse. Due to the relative phase conjugation between A* and C, the evolution of the absorption and emission have the opposite sign.

again all have the same phase, and so they are said to have rephased. The response due to rephasing is called a photon echo.

Other detection frequencies correspond to different pulse sequences. The most common are the non-rephasing sequence and the double-quantum sequence. These pulse sequences are used in Chapters VI and IX. The non-rephasing sequence, which can be measured at the frequency $\omega_A - \omega_B + \omega_C - \omega_{\mathrm{LO}}$, is useful for measuring homogeneously broadened systems. Rephasing and non-rephasing measurement of a homogeneously broadened system will be identical. In Chapter VI we use the sequence to better isolate individual quantum dots (single oscillators confined in a semiconductor) from the ensemble of weakly excited quantum dots that contribute to a background signal in rephasing spectra. We use double-quantum MDCS, where both the first and second pulses have the same conjugation, to selectively measuring many-body interactions. This signal is measured at frequency $\omega_A + \omega_B - \omega_C - \omega_{\mathrm{LO}}$, where care must be taken to ensure the signal phase is still preserved. We use this

to measure weak interactions between quantum dots in Chapter VI. Higher-order nonlinear signals can be used to probe high-order correlations [30]. Though not presented in this thesis, we can measure N-quantum MDCS just by changing the detection frequency [31].

There are now numerous groups that use collinear MDCS to measure semiconductor, atomic, molecular and biological systems. The majority of these excite the sample with four pulses and measure incoherent population-based signals. The population signals are typically photoluminescence (fluorescence) [13, 15, 32] or photocurrent [26]. Recently mass spectrometry was used as the population signal for this type of measurement [33]. Our measurements here are all coherently detected, so the analysis of these results is more similar to that of conventional wave-vector selection spectroscopies. The most similar spectroscopy to this is the coherently detected four-wave mixing (FWM) techniques of W. Langbein and J. Kasprzak [34, 35].

2.8 Spatial Chirp Correction

To frequency shift an optical beam, we diffract that beam off an acoustic wave travelling through an AOM. We change the frequency shift of the optical beam by changing the frequency of the acoustic wave. The light is diffracted at angle θ, which is given by Bragg's law:

$$\sin\theta = \frac{n\lambda}{2d} = \frac{n\lambda f_{\mathrm{mod}}}{2v_s}, \tag{2.9}$$

where λ is the wavelength of light, f_{mod} is the modulation frequency, and v_s is the speed of sound in the crystal.

A common problem with MDCS techniques that use frequency tagging is spatial chirp. The beams are all frequency tagged with AOMs, but the diffraction angle of the frequency shifted beam is wavelength dependent. Many spectroscopists use lasers with very broad bandwidths to excite a wide range of resonances. Nobody wants to describe

the drawbacks of their techniques in writing, but in talks and private conversations the problem of spatial chirp is commonly mentioned. The typical solution is to spatially filter the beam after the AOM, which is a fairly lossy correction. There is an easier way.

In atomic physics there is a similar problem that AOM user face. While the beams are typically monochromatic, atomic physics use AOMs to shift the center frequency by up to tens of MHz. Since the diffraction angle also depends on the modulation frequency, frequency tuning the beam would deflect it and change the experiment alignment. The solution was to double pass the AOMs [36]. The deflected beam is reflected so that is passes back through the AOM, and the second deflection corrects the spatial chirp. We have implemented this in our setup and found that it seems to completely correct the spatial chirp issue for a broadband beam having over 60 nm of bandwidth.

2.9 Improved Reference Technique

One problem with the standard phase correction scheme is that it requires a CW reference laser with a center wavelength that is very near that of the signal. While the pulsed lasers commonly used for MDCS to measure a variety of samples are very tunable, the CW lasers we use are not. Almost every time we change samples, we have had to change the laser diode in the external-cavity diode laser (ECDL). To solve this problem, we have invented a method for generating a demodulation frequency (that samples path length fluctuations of the interferometer) and maintaining a record of the path length difference using a reference laser having an arbitrary frequency with respect to the light measured. This is a method that I have developed with Chris L. Smallwood and Steven T. Cundiff.

For the improved reference technique we use a field-programmable gate array (FPGA) to calculate the path length difference. The optical phase difference between

the two paths is determined by measuring the intensity modulation of the interference between the reference laser following each path and subtracting this from the intended modulation frequency. This optical phase is equal to the modulus of the path length difference and the reference laser wavelength, multiplied by 2π. The overall phase (equal to the ratio of the path length difference and the reference laser wavelength, multiplied by 2π) would be ambiguous for path differences separated by greater than the reference laser wavelength, except that we unwrap the phase using the FPGA. So long as the evolution of the interference is well sampled (there must be at least four data points per full modulation), the phase can be unwrapped in real-time. Calculation of the path length difference is then the product of the overall phase and the reference laser wavelength.

The demodulation frequency of the signal is also calculated by the FPGA in real time. The overall phase of the signal is added to the intended modulation frequency, thus generating the corrected demodulation frequency.

In order to measure an optical signal with phase sensitivity it is necessary to sample all the path length fluctuations with a reference laser. In practice, we do this by subtracting measured phase fluctuations of a reference laser from the the signal phase. In Table 2.1, we show how a detuned reference affects the measured signal. We generate this table for a signal having an optical wavelength of 800 nm and path length fluctuations within the time of measurement being δx =500 nm. The second column, $\delta\phi_{sig} - \delta\phi_{ref}$, indicates how well the reference fluctuations cancel the signal fluctuations. The amplitude of the measured signal is degraded by these phase fluctuations by a factor of $1 - \delta A = cos(\delta\phi_{sig} - \delta\phi_{ref})$. A reference having the same wavelength as the signal could perfectly cancel fluctuations. A slightly detuned reference having a wavelength of 810 nm would nearly cancel all phase fluctuations of the reference and has negligible effect on the measured amplitude. A very detuned reference having a wavelength of 1064 nm would very badly correct phase fluctuations

and would underestimate the signal by over 40%.

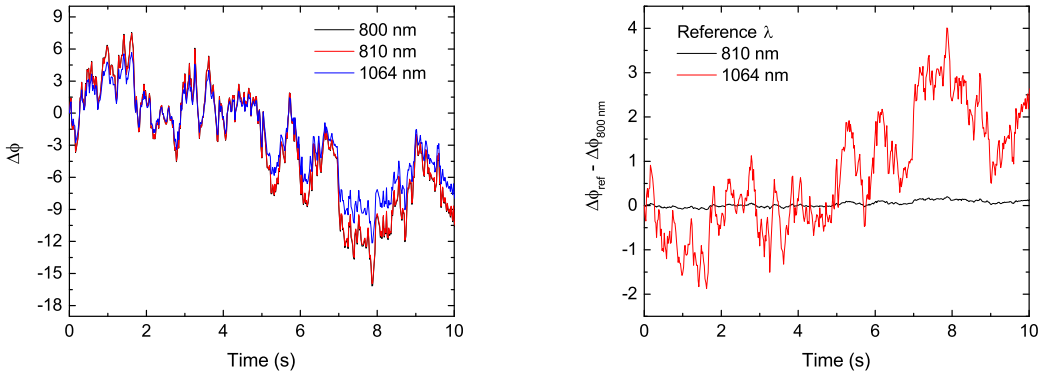
Table 2.1: Fluctuation correction of a signal having a wavelength of 800 nm by a reference laser having various wavelengths. Large-detuned reference wavelengths badly cancel the phase, and this leads to large degradation of the measured signal amplitude.

λ_{ref} (nm)	$\delta\phi_{sig} - \delta\phi_{ref}$ (rad)	δA (unity normalization)
800	0	0
810	0.048	0.001
1064	0.974	0.438

In Figure 2.8(a), phase fluctuations between two arms of an interferometer are plotted for the same simulated path length fluctuations with time. These are plotted for light having a wavelength of 800, 810, and 1064 nm. In Figure 2.8(b) the difference between these curves illustrates how incorrect a phase measurement could be after a few seconds. This means that for a reference at 810 nm cancelling fluctuations of an 800 nm signal, one would always measure a nearly correct phase. For a reference at 1064 nm, the phase could easily be measured incorrectly by over π (which corresponds to measuring a signal with the exact opposite sign of what is should have). It is thus apparent that one cannot do phase-resolved spectroscopy using a far-detuned reference and the standard phase cancellation scheme. Really the phase difference should not vary by more than 0.15 rad or else the data could be misinterpreted (using this rule we could define far detuned as greater than 30 nm at a center wavelength of 800 nm, or 60 meV in energy units).

Instead, it is necessary to consider the wavelength difference and scale the phase fluctuations with the reference wavelength. One could naively think of just multiplying the phase by the ratio of the reference and signal wavelengths. The problem is that one can only measure phase within a range of 2π, at which point the phase wraps. In Figure 2.9(a) we plot the measured relative phase between two arms of an

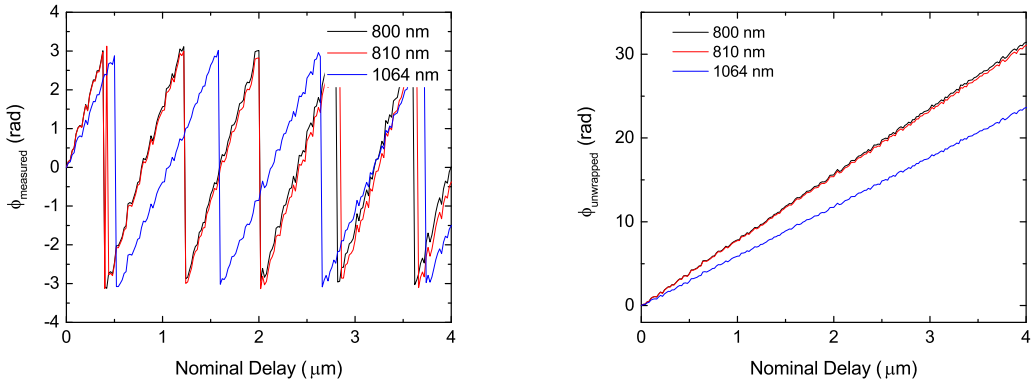
interferometer in which one arm is intentionally delayed with respect to the other. In the figure one can see the phase evolving as the delay is scanned, and the phase evolves differently for different wavelengths. Due to the phase wrapping, there is no simple way of correcting the phase without some record of the number of wraps between the delay and zero delay. By unwrapping the phase, shown in Figure 2.9(b), we can multiply the actual reference phase by the ratio of the reference and signal wavelengths. This unwrapping can be done by recording the number of phase wraps in an FPGA. This new calculated phase for any optimized wavelength is applied to the modulation frequency so that the modulated signal can be measured.



(a) (b)

Figure 2.8: Relative phase fluctuations as a function of time for various wavelengths of light in an interferometer. (a) As the relative path length fluctuates with a standard deviation of 1 micron, the relative phases fluctuate as well. (b) The measured phase of the signal will be wrong if the wavelength of the reference does not equal the wavelength of the signal. The amount that the phase measurement is wrong is plotted here for a signal at 800 nm and a reference at 810 and 1064 nm.

One other method we have considered for correcting the signal phase with a far-detuned reference laser is the use of adaptive sampling. This is a different calculation that can be performed by an FPGA to also cancel path length fluctuations. For this method, rather than cancelling the phase of the signal, the sampling rate of the signal is varied with the path length fluctuations. Rather than sampling the signal every 10 microseconds, the signal is sampled every time the reference phase



(a) (b)

Figure 2.9: Here we demonstrate the necessity of unwrapping the phase of the measured reference in order to calculated a corrected reference for an arbitrary wavelength. (a) For delays greater than the wavelength of light, the wrapped phase of a signal at one wavelength cannot be related to a reference at a far detuned wavelength. (b) If the phase is unwrapped, the fluctuations can be corrected exactly by multiplying the reference phase by a ratio of wavelengths.

is incremented 0.001 radians (about equal to 1 nm). This is more similar to the fringe-counting scheme used in Fourier-transform infrared (FTIR), except frequency modulation allows us to take more steps. We would accomplish this scheme by locking the FPGA clock to the reference signal. This method would really only work for moving delays (otherwise the sampling rate is 0), and it would require fluctuations in the negative direction be treated carefully.

This fluctuation correction scheme is useful for generating a lock-in reference for collinear spectroscopies and for actively correcting path length fluctuation in a frequency-modulated interferometer. Though we apply it to MDCS, it would be useful in any interferometer. The scheme also highlight a major feature of frequency-modulated interferometers, which are not commonly used in one-dimensional Fourier-transform spectrometers: frequency modulation allows the use of an equal or long wavelength reference with respect to the signal. This is because the zero-crossings of the reference are no longer tied entirely to the relative path length as in the FTIR standard. Therefore modulation would allow much greater versatility and higher

sensitivity of interferometer sampling in applications including FTIR.

2.10 Continuous Scanning

Typically we delay pulses in time using discrete steps. We measure a range of delays in t and τ defined by where the signal is greater than the noise. We scan over that range using equidistant steps in t and τ that are defined by the required measurement bandwidth. A measurement at each point is as follows: move stages that control t and τ delays to the point of interest, wait for the stages to settle and for the lock-in amplifier to converge on the correct values, and acquire and save the measurements from the lock-in. Depending on the lock-in time constant and averaging, this can take between 400-1500 ms. Since these are two-dimensional measurements, the total acquisition time scales with the number of steps along one dimension squared. High-resolution measurements of large-bandwidth systems can therefore take a very long time. Though we have managed 1-5 minute typical acquisition times of the multidimensional spectra in this thesis, doing so has required restricting the laser bandwidth in some experiments.

We have largely solved the issue by continuous scanning. For continuous scanning we continuously move one delay stage corresponding to one time axis and continuously record measurements from the lock-in amplifier. The other time axis is still scanned discretely. We have implemented a rudimentary version in the lab that drastically improves acquisition times and increases the detection bandwidth by over an order of magnitude without sacrificing the signal-to-noise ratio. We currently determine the stage position by assuming a constantly moving stage, which is a false assumption that leads to some linewidth broadening. However, when continuous scanning is paired with the path-length monitor described in the previous section it will be immune to any inconsistencies in the delay stage movement.

For continuous scanning there a few things we must consider when selecting a

stage-delay speed and lock-in time constant TC. Since the final stage of a lock-in amplifier is functionally a bandpass filter, it integrates the signal from the detector over a time that depends on the filter bandwidth and slope. For a slope of 12 dB/octave, the signal will stabilize to with 99% of the correct value within 7 TC [37]. TC therefore roughly sets our new effective step size since data points sampled within that step size are significantly correlated. Using this time constant and the stage velocity (v_s) we determine that the effective optical bandwidth of the measurement is $\sim \left(\frac{2v_s}{c} \cdot TC\right)^{-1}$, where v_s is multiplied by two because the delay stage is double passed. The distinction now is that this time axis is not undersampled, so this bandwidth refers to the fastest varying signal with respect to the reference. If the reference technique described in the previous section is used, the effective reference frequency should be placed near the signal frequency to minimize the required bandwidth. A stage velocity of 1 mm/s and a time constant of 10 ms corresponds to an optical bandwidth of roughly 15 THz $\sim$ 60 meV. For a very flat detector response across the entire bandwidth, one should only use $1/7^{th}$ of this bandwidth for considering 99% convergence of each signal within the time constant.

2.11 Treatment of Linear Dispersion

Broadband light will be significantly temporally chirped after transmitting through an AOMs and glass in the setup. It is therefore standard to correct for dispersion in all the excitation beams using a prism pair, a pulse shaper, or chirp-compensating mirrors [38]. Though the chirp of the beams exciting the sample can be completely corrected, the chirp on the LO beam will most likely be different because it passes through a different set of optics. Though one could independently correct this chirp, there is a much simpler way to treat the relative chirp between the LO and excitation beams.

Using two lock-in amplifiers we simultaneously measure both the nonlinear MDCS

signal and the linear inference of the LO and the third pulse. Both of these interference terms modulate the detector output, so it is really just a matter of generating the reference frequencies for the two lock-ins. The linear interference of the LO and third pulse corresponds to a cross-correlation of those two beams. Fourier transforming the cross correlation between LO and the third pulse reveals the relative linear spectral phase of the two beams. If there were no relative linear spectral phase between the LO and third pulse, the interference spectrum between them would be purely real. By measuring and subtracting the relative spectral phase, it is thus possible to rotate the linear spectrum so that it is purely real. This same rotation is applied to the nonlinear signal to determine its spectral phase in the absence of any linear phase rotations. This method is used to subtract dispersion from the differing thicknesses of glass and other transmissive elements between the LO and excitation beams. It also subtracts the linear phase applied to the excitation beams and nonlinear signal by the sample. The linear phase of a strong resonance can make a nonlinear signal look dispersive when, in fact, it is purely absorptive. It is possible to automatically subtract this linear phase by passing the LO to the sample along with the other beams, but doing so can further modify the measurement of the nonlinear response through $\chi^{(5)}$ terms that become significant [39]. We therefore route the LO around the sample and subtract the measured relative linear phase. In Fig. 2.10 we plot the real part of the MDCS measurement measured with broadband light. The sample is a heterostructure of monolayer transition metal dichalcogenides that have two coupled resonances and some other features. Most notable here we see in the left plot that the relative chirp between the signal and LO is so significant that the low energy features have visibly lower frequency oscillations than the higher energy features. In Fig. 2.10 we demonstrate that we can correct this chirp by correcting the phase with a simultaneous linear measurement (not plotted).

In summary, the simplest method for dealing with temporal chirp is to dispersion

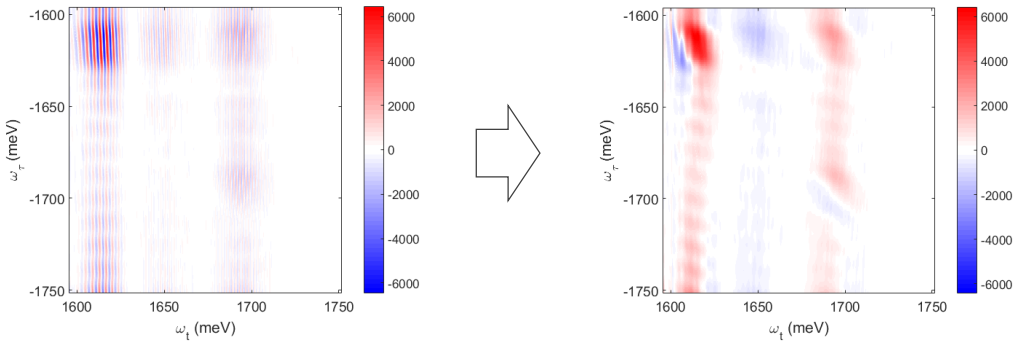


Figure 2.10: Real part of the MDCS measurement of two coupled resonances excited using broadband light. The relative chirp between the LO and excitation beams was significant such that in the left plot the spectral oscillations of the low energy features are visibly slower than those of the high energy features. On the right we plot the same measurement after implementing the chirp correction.

compensate all beams to excite the sample with nearly bandwidth-limited pulses. This may not mean that the LO is bandwidth limited, but it will ensure that a linear phase does not affect the nonlinear signal. Interestingly, it is not essential to correct the phase of the excitation pulses. By measuring multidimensional spectra as a function of the delay between pulses B and C it is possible to project the spectra so as to remove distortions due to pulse chirp [40].

2.12 Fitting Multidimensional Spectra

We quantify the linewidths measured with MDCS by fitting the spectra with simple models. We model a system excited by ultrafast pulses in the time domain. In the Markovian approximation this results in the signal:

$$s(t,\,\tau) = s_0 e^{(\gamma(t+\tau)+i\omega_0(t-\tau)+\sigma^2(t-\tau)^2/2)}\Theta(t)\Theta(\tau), \tag{2.10}$$

where γ is the homogeneous linewidth of all the oscillators, ω_0 is the center frequency, and σ is the inhomogeneous linewidth. The τ and t times correspond to the delays between the first and second pulses and the emission time after the third pulse,

respectively. We Fourier transform this signal with respect to τ and t to generate two-dimensional spectra.

The two-dimensional frequency spectra are measurements of the FWM amplitude as a function of absorption (ω_τ) and emission (ω_t) energies. The cross-diagonal linewidth roughly corresponds to the homogeneous linewidth, and the diagonal lineshape roughly corresponds to the inhomogeneous distribution. We see that this treatment of inhomogeneity is identical to convolving a homogeneous oscillator with a Gaussian distribution in the frequency domain. It is thus best to think of a two-dimensional spectrum as a sum of homogeneous oscillators aligned along the diagonal. Though we have simple analytic solutions to the optical Bloch equations (OBEs) for homogeneous dephasing and a Gaussian distribution of oscillators, we can numerically fit experimental data with arbitrary distributions and linewidths. Siemens *et al.* details the simplest treatment and the fitting procedures we typically use to simultaneously fit the codependent diagonal and cross-diagonal slices [41]. This paper also includes depictions of the star shape that a homogeneous feature has in a two-dimensional plot. We emphasize here that the cross-diagonal lineshapes of inhomogeneously-broadened and homogeneously-broadened systems are very different. In the limit of high inhomogeneity the lineshape is the square root of a Lorentzian, which results from measuring the amplitude of the signal. The full width at half maximum of this signal is therefore $\sqrt{3} \times 2\gamma$. For no inhomogeneity the cross-diagonal lineshape is a Lorentzian, which has a full width at half max of 2γ [41]. Ignoring inhomogeneous broadening and using the wrong fit function when determining a sample's homogeneous linewidth (or dephasing times in photon echo FWM experiments) can thus significantly skew the measurement, up to a factor of $\sqrt{3}$.

We can also analytically fit full multidimensional spectra with solutions to the OBEs derived by Bell *et al.* [3]. For lineshapes that are sufficiently broad to require considering the laser bandwidth, we fit spectra with analytical solutions that include

finite pulse effects [42]. In Fig. 2.11 we show a characteristic two-dimensional plot and a full fit of the plot.

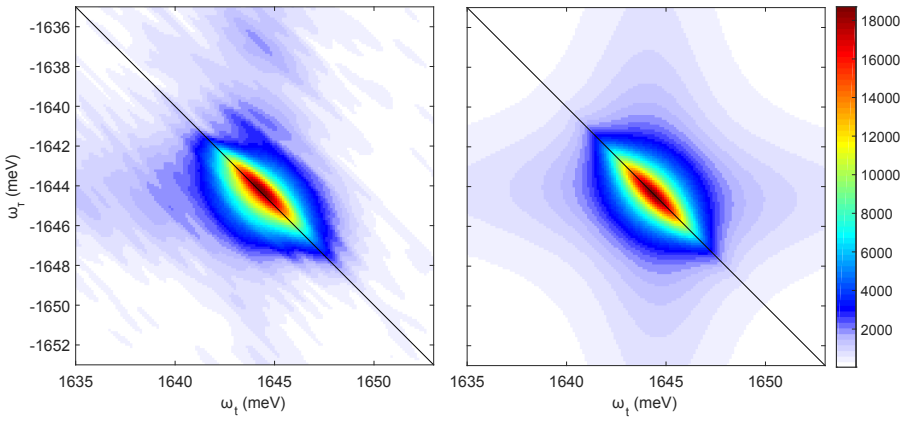


Figure 2.11: Characteristic MDCS measurement of an encapsulated $MoSe_2$ sample. The linecut along the diagonal (where $\omega_\tau = -\omega_t$) corresponds to the inhomogeneous distribution and the cross-diagonal corresponds to the homogeneous lineshape.

All of the analytical fit functions described above assume the dephasing rate across the the inhomogeneous distribution is homogeneous. In quantum wells it has been demonstrated that this is not the case [29, 43]. For instance, if the source of the inhomogeneity is localization one would expect the lower energy resonances to have lower dephasing rates. We have thus derived a solution to the OBEs for an inhomogeneous dephasing rate with a linear dependence on energy. We begin by defining the dephasing rate: $\gamma = \gamma_0 + \gamma_\delta \omega_0$, where γ_0 is the homogeneous linewidth at the center of the resonance and γ_δ is the slope of the linewidth change with energy. We can rewrite the complex resonance frequency so that the frequency dependent term now has an imaginary component.

$$\omega_0 - i\gamma \rightarrow \omega_0(1 - i\gamma_\delta) - i\gamma_0. \tag{2.11}$$

Following the steps in [3], it is straightforward to show that the rephasing signal is

$$
\begin{aligned}
S_R(\omega_t,\,\omega_\tau) \propto & \frac{1}{2\sigma(1-i\gamma_\delta)(2\gamma - i(\omega_t+\omega_\tau))} \times \\
& \left(e^{\frac{(\gamma - i(\omega_t-\omega_0))^2}{2\sigma^2(1-i\gamma_\delta)^2}} \operatorname{Erfc}\left[\frac{\gamma - i(\omega_t-\omega_0)}{\sqrt{2}\sigma(1-i\gamma_\delta)}\right] \right. \\
& \left. + e^{\frac{(\gamma - i(\omega_\tau+\omega_0))^2}{2\sigma^2(1-i\gamma_\delta)^2}} \operatorname{Erfc}\left[\frac{\gamma - i(\omega_\tau+\omega_0)}{\sqrt{2}\sigma(1-i\gamma_\delta)}\right] \right).
\end{aligned}
\tag{2.12}
$$

This result is nearly identical to the result for homogeneous dephasing except that γ now has a spectral dependence and the inhomogeneous linewidth now has an imaginary component: $\sigma \to \sigma(1 - i\gamma_\delta)$.

CHAPTER III

Excitons in Semiconductor Nanostructures

The interaction of semiconductors and light is of practical and fundamental importance. Laser diodes (found in printers, light shows, range sensors in autonomous cars, etc.), light emitting diodes (LEDs), and displays are all based on light emission by semiconductor nanostructures. Solar cells, solar panel windows, digital cameras, and photodetectors are all based on light absorption and conversion to electrical current by semiconductor nanostructures. In order to improve the efficiency, functionality, and production of all of these devices, it is essential to understand the fundamental interactions of excitations created in semiconductor nanostructures by light. It is important to understand the fundamental loss mechanisms that impede device efficiency, the many-body interactions that dominate the optical emission of highly-excited semiconductors (e.g., lasers) [10], and the effects of inhomogeneity on the functionality of the semiconductor.

Many-body interactions are known to dominate the nonlinear optical response of semiconductors [44, 45]. In these material, 10^{22}–10^{23} particles cm^{-3} are interacting via the infinite-range Coulomb force, which is an incredibly complex theoretical problem to solve. However, decades of study have developed an advanced theoretical understanding of the linear properties of semiconductors [7]. The introduction of quasiparticles to describe complex interactions has tremendously simplified the treatment of many-body systems. Beyond these treatments, the nonlinear optical

responses of semiconductors are fundamentally different from atoms. As a relatively well-understood condensed system with residual unexplained effects, semiconductors and semiconductor nanostructures are ideal systems for studying many-body physics.

3.1 Bulk Semiconductors and Band Structure

Materials can be classified as conductors, semiconductors, or insulators depending on their ability to conduct electricity. Metals are very good conductors, insulators are very poor conductors, and semiconductors can transition between the two. The origin of the material conductivity and the qualitative distinction between these classifications is better understood by considering the electronic band structure.

In a periodic lattice, the single-particle eigenstates are Bloch waves. The translation symmetry of a periodic Hamiltonian implies that Bloch waves are eigenstates of the translation operator. For a translation of the lattice period a_0, the operator is $\hat{T}(a_0) = \exp(-ia_0\hat{k})$. $\hat{k}$ here is the momentum operator, which is the crystal momentum. We find that the allowed energies of an electron in a crystal are therefore defined as a function of the crystal momentum $\mathbf{k}$. The energies $E(\mathbf{k})$, which describe the electronic band structure, are the energies of the electronic states that an electron can occupy. Along a single k axis, the band is only defined in the range from 0 to π/a_0. The dispersion relation defined for $\pi/a_0 < |k| < 2\pi/a_0$ thus defines the second energy band, $2\pi/a_0 < |k| < 3\pi/a_0$ defines the third, and so on [46].

The electronic bands describe the available states of the system, and electrons fill these states from the lowest energies up. The valence band describes the highest filled energy level at T = 0 K, and the conduction band describes the first unfilled level above the valence band. In metals, there is no gap between these bands, and so electrons are easily scattered into unoccupied states with the application of an external field. In insulators the gap is very large so that chemical doping, application of strong electric fields, and irradiation with light do not effect the occupation of the

conduction band. These band gaps, E_g, are typically on the order of 10 eV, though the actual value is somewhat dependent on the application. In a semiconductor there is a small gap between the valence and conduction bands. In silicon the band gap at 0 K (300 K) is 1.17 eV (1.1 eV), and in gallium arsenide (GaAs) the gap is 1.52 eV (1.4 eV) [47]. The room temperature band gap of silicon corresponds to a photon having the wavelength 1130 nm.

Valence band maxima and conduction band minima typically lie at high-symmetry points where the degeneracy of adjacent bands leads to an avoided crossing [48]. These points do not always exist at the same high-symmetry point, though. Materials in which the band points nearest in energy are at the same crystal momentum have what is called a direct band gap. Materials for which those points are at different crystal momenta have an indirect gap. This distinction becomes interesting when considering optical excitation of an electron from the valence band to the conduction band. Photons have such little momentum, that a photon that is resonant with the band gap of an indirect semiconductor cannot directly excite an electron. The absorption coefficient for indirect band gap semiconductors is very low because the process requires the simultaneous absorption of a photon and a lattice vibration, called a phonon, to have sufficient energy and momentum for the electron transition.

Though band structures are typically quite complicated, the energies near the fundamental band gap can typically be approximated as depending quadratically on the crystal momentum. The energy of band λ near the high-symmetry points can be written $E^\lambda(\mathbf{k}) = E_g^\lambda + \frac{\hbar^2\mathbf{k}^2}{2m_{\text{eff}}}$, where m_{eff} is the effective mass that can be used to describe the system with a single-particle Hamiltonian. In the two-band approximation for a direct-gap semiconductor, useful for processes that can be described with just two bands, the conduction and valence bands are described with effective electron (m_e) and hole (m_h) masses. Since the dispersion of the valence band is negative, the quasiparticle hole has a negative effective mass. Describing semiconductors excita-

tions with quasparticle electrons and holes is sufficient for solving many problems in semiconductors, but this neglects the interactions between particles.

3.2 Excitons

The Coulomb interaction between electrons and holes introduces a new quasiparticle called an exciton. The exciton is a bound state between the quasiparticle electron described by the conduction band and a hole described by the valence band. The wave function of an exciton therefore has components of both the Bloch wave function and the hydrogenic wave function describing the binding between two particles having opposite charge [46]. From this we realize that the exciton energy is equal to $E_g + E_0$, where E_0 is the energy of the hydrogenic part of the wavefunction, which is negative for bound states. We also realize that, similar to electrons, excitons are composed of states that are delocalized over the entire crystal [49, 50].

In the dipole approximation the macroscopic polarization induced by a field $E(\omega)$ can be calculated by summing over the polarizations of all bound and unbound states in the excitonic wave function $\psi(\mathbf{r})$. Focusing on direct-gap excitons, the linear semiconductor susceptibility is then:

$$\chi(\omega) = \frac{|d_{cv}|^2}{\epsilon_0} \sum_n \frac{|\psi_n(\mathbf{r}=0)|^2}{E_g + E_n - \hbar\omega - i\gamma}, \tag{3.1}$$

where d_{cv} is the transition dipole moment and γ is a phenomenological term describing dephasing. This is the famous Elliott formula that we use to describe the linear interaction of light with semiconductors. The formula introduces exciton resonances below the band gap. It is important to note that the presence of excitonic resonances does not immediately imply the existence of exciton population because the ground-state system has no electron-hole pairs [51]. This can further be understood by realizing that true photon absorption (and thus creation of exciton population)

requires a second-order process that scatters energy from the exciton mode into the crystal lattice [49].

3.3 Characteristics that Distinguish Semiconductor Samples

In this thesis we study a variety of samples, and we seem to apply a different technique to each one. The reason is that the properties of semiconductor samples differ widely, and the characteristics are dependent on the semiconductor material, the sample confinement, the effect of the device structure on the impinging light, and the sample conditions. We therefore find that different samples emphasize different features of excitons. In this section we introduce a few of the characteristics of semiconductor samples that distinguish the samples measured here. In the next section we compare the various samples directly.

3.3.1 Inhomogeneity

Inhomogeneity of the center energies of an ensemble of resonances broadens optically measured linewidths. There are numerous sources of inhomogeneity, and many of these sources affect the sample response much more than just broadening the measured optical linewidth. Nonlinear optical techniques, and in particular MDCS, enable the measurement of intrinsic sample linewidths in the presence of inhomogeneity. They also enable measurement of the dynamics of and coupling between energetically distinct states.

Confinement increases the band gap of a semiconductor. This is easily understand by considering the single particle wave function in a finite square well. The wave function extends into the barrier, which increases the ground-state energy levels. In an infinite square well, the particle ground-state energy depends inversely on the square of the well length. This confinement effect results in an increase of the semiconductor band gap. The confinement of excitons also affects their binding energies. The binding

energy decrease is a smaller effect (with the opposite sign) than the band gap increase on the transition energy of the exciton.

A quantum well (QW) is a structure that confines an exciton in one dimension, and a quantum dot (QD) is a structure that confines an exciton in three dimensions. Inhomogeneity of the transition energies in these samples is due to spatial fluctuations of the confinement potential. In QWs the well width varies slightly over the sample. The wider regions of the QW have lower energies and can trap excitons [52]. Even narrow inhomogeneous distributions can therefore describe excitons with very different wave functions and properties.

There are other sources of inhomogeneity. The exciton energy can depend on the electrostatic environment. This means that trapped charges and surface states that locally modify the crystal potential can influence the exciton center energies. Inhomogeneous strain can also modify the local band structure and influence the exciton energies [53]. For strongly confined excitons that have wavefunctions extending outside their crystal of origin, dielectric inhomogeneity of the crystal enivironment can also introduce inhomogeneous broadening.

3.3.2 Decay/dephasing Processes

The dynamics that occur after optical excitation of an exciton are typically relaxation, dephasing, and spatial diffusion. Though we explore spatial dynamics in Chapter V, we will focus on relaxation and dephasing here. We also focus on processes that contribute to homogeneous decay and dephasing. Inhomogeneity will also dephase the macroscopic polarization, but we can extract the homogeneous linewidths with MDCS.

Radiative decay processes, which have the rate Γ_{rad}, are possibly the most fundamental decay process in semiconductors. Γ_{rad} corresponds to the strength of coupling between the exciton and light, and it gives the exciton some fundamental linewidth. It

is often related to the timescale of exciton recombination by the spontaneous emission of a photon for excitons, but this only applies to optically active excitons. Optically active excitons are those excitons that have a wavevector within the light dispersion line ($k > \frac{E}{\hbar c}$). There are ways to enhance the radiative linewidth, discussed in the following subsection, but this linewidth is difficult to measure otherwise.

The homogeneous linewidth γ is also broadened by incoherent processes. Instead of radiatively decay to the ground state, excitons can scatter into another excited state, which is a source of spectral diffusion [54]. Shown in Fig. 3.1, these processes can occur with absorption or emission of a phonon. They can also occur with the simultaneous excitation of another exciton through an Auger process. Whether scattering into another band, another valley of the same band, or into a localized state, the process will increase the decay rate of the optically active exciton (by the rate of spectral diffusion Γ_{SD}). These other states are often dark because transition out of them would require a photon with an impossibly high momentum. Transitioning out of dark states is therefore a typically slow process that requires an additional interaction with phonons to either become optically active or to relax back to the ground state. Relaxation of dark states by phonon emission is accompanied by sample heating. Throughout this thesis we explore some of the decay mechanisms that can affect exciton linewidths.

Scattering with a phonon or another exciton does not always cause the exciton to change its energy. Elastic scattering processes can also affect the exciton linewidth, though they do so by dephasing the exciton with respect to the excited ensemble. Dephasing will shorten the coherence time of a transition without actually inducing decay of the transition. We call this linewidth contribution pure dephasing, or γ^*.

The various decay and dephasing processes are revealed in various spectroscopies.

- The macroscopic polarization decay, which corresponds to the inverse of the transition linewidth, is sensitive to all broadening mechanisms including de-

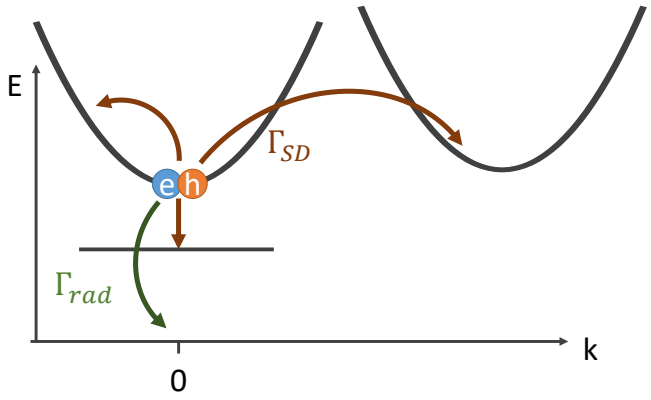


Figure 3.1: Excitons created at $k = 0$ can radiatively decay back into the ground state with a rate Γ_{rad}, or they can scatter into other excited states. We draw a hypothetical exciton band structure containing two relevant valleys. In direct gap semiconductor, the valley at $k = 0$ is at a lower energy than other valley. The exciton can scatter by absorption or emission of a phonon into a higher energy state of the same valley, into another valley, or into lower energy localized states if they are available.

phasing by inhomogeneity.

- The photon echo corresponds to the measure of the homogeneous decay. Using MDCS we can resolve the homogeneous decay as a function of energy in the inhomogeneous distribution. The linewidth γ corresponds to a coherence time T_2.

- Transient absorption measures population decay times. Spectrally resolved transient absorption can distinguish differences in population decay for different populations in the inhomogeneous distribution. Typically one will measure two timescales of population decay: the fast timescale corresponds to radiative decay and spectral diffusion into dark states. The slow timescale corresponds to the decay of dark states either by scattering back into optically active states or through phonon emission.

- Time-resolved photoluminescence (PL) is another probe of the exciton population, but it is only sensitive to the optically active population. The initial fast

timescale of PL measures similar processes to transient absorption, though this is typically very difficult to measure. The slow timescale of PL, on the other hand, is only sensitive to decay of dark states via optically active excitons.

We have listed the quantities that describe the various decay times, but let's summarize them here. The optical population decay time, T_1, corresponds to the fast decay time of an optically excited transition. It is inversely related to the decay rate, $\Gamma = 1/T_1$. The decay rate includes both radiative decay and spectral diffusion, so $\Gamma = \Gamma_{rad} + \Gamma_{SD}$, where Γ_{SD} is the rate of spectral diffusion by inelastic scattering. The slow population decay time corresponds to the rate of spectral diffusion back into the optically active states and non-radiative decay rate. This rate is generally very low. The homogeneous linewidth and the coherence time are inversely related: $\gamma = 1/T_2$. For a system without pure dephasing, the decay time and the coherence time are typically related by a factor of two, where the optical population decay is twice as fast as the coherence decay. However, the homogeneous linewidth is also affected by pure dephasing. Therefore $\gamma = \Gamma/2 + \gamma^*$. Since we are measuring homogeneous linewidths, we will sometimes lump all inelastic scattering and pure dephasing into a single term defining the scattering rate $\gamma_{sca} = \Gamma_{SD}/2 + \gamma^*$.

3.3.3 Exciton-photon Coupling Strength

Exciton-polaritons are quasiparticles that arise from the strong coupling of excitons and photons. There are two types of exciton polaritons. 1) The exciton resonances measured in cryogenically cooled bulk semiconductors are evidence of exciton-polaritons. Without scattering (γ_{sca}), the interaction of light with a bulk material is not accompanied by actual absorption. A photon that transmits through a semiconductor and is resonant with an exciton transition will interact with a strength given by the radiative linewidth of the transition Γ_{rad} [55]. If the interaction strength is greater than the scattering rate, the light-matter interaction is strong. So without

scattering, the transmitted photon becomes a polariton in the bulk semiconductor. Upon exiting the crystal, the photon state is recovered. The phase of light passing through the crystal will be modified by the crystal, but no energy is lost.

2) In semiconductor nanostructures, for instance GaAs QWs, the scattering rate typically dominates the radiative decay rate even at low temperatures. To introduce strong light-matter coupling, the interaction of light with QWs can instead by enhanced by increasing the coherent interaction length of the light and matter, thereby enhancing the effective radiative linewidth of the QW. This enhancement is accomplished by creating a system in which light interacts with QWs spaced by $\lambda_{ex}/2$, where λ_{ex} is the wavelength of the exciton transition, so that the induced macroscopic polarizations constructively interfere in the forward and backward directions. These many-QW samples with Bragg periodicity have been grown, and the exciton-polariton interaction dominates the optical response [56]. Exciton-polaritons are now more commonly generated by placing a QW in an optical cavity that is nearly resonant with the exciton transition [57]. The radiative linewidth enhancement can be similarly understood to the Bragg-spaced many-QW samples as the the light interacts with the sample many times before exiting the cavity. Normal-mode coupling is an additional feature of these types of exciton polaritons, where the cavity mode strongly couples to the exciton mode introducing upper and lower polariton modes.

3.3.4 Motional Narrowing of Microcavity Polaritons

Returning to the discussion of inhomogeneity, inhomogeneous materials placed in a cavity can exhibit motional narrowing. The cavity introduces a significantly increased coherence length that averages over the spatial inhomogeneity of the material. The distribution of the spatially averaged energies is typically much less than the homogeneous linewidth of cavity-exciton polaritions [58]. Polariton samples are therefore among the few truly homogeneous semiconductor nanostructures.

3.4 Comparison of Samples Measured Here

GaAs QWs are very well understood semiconductor nanostructures. These samples are epitaxially grown with a typical width of 10 nm and barriers made of $Al_{0.3}Ga_{0.7}As$. The binding energy of excitons in these structures is approximately 10 meV, depending on the QW size. These structures have homogeneous and inhomogeneous linewidths that are roughly 0.1 meV and 1 meV at low temperature, respectively [1]. The spatially inhomogeneous potential localizes excitons and introduces a high number of dark states. The scattering rate is correspondingly much larger than the radiative decay rate [59].

Interfacial quantum dots are formed by well-width fluctuations of a thin GaAs quantum well [60, 61, 62]. The nominal width is roughly 15 monolayers (4.2 nm), which is narrow enough such that monolayer width fluctuations correspond to roughly 10 meV shifts of the exciton resonance. The growth of the top barrier is delayed by tens of seconds to give monolayer width fluctuations of the quantum well time to coalesce into islands. These islands, with a width of 16 monolayers, form the interfacial quantum dots, which are bound by 10 meV. The QDs have a lateral size of approximately 36 nm [63]. By measuring individual dots we can measure resonances that are homogeneously broadened. The localized exciton linewidths are also radiatively limited at low temperature because they cannot relax into any lower energy states.

Zero-dimensional polaritons are studied in a 7.5 μm device that confines cavity-exciton polaritons in three dimensions. The devices are made by growing GaAs QWs at the anti-nodes of a cavity made by a high reflector and a sub-wavelength grating (SWG). The QW confines the polariton in the z-direction, and the finite size of the SWG confines the polariton in the other two dimensions. Strong interaction between the exciton and photon give the polariton states a very low effective mass. Though the polaritons are confined, the coherence length of these polaritons is very large. Some of the higher-order wave functions extend across the device, and in

Chapter IX we measure coherent coupling over multiple micron ranges.

Monolayer transition metal dichalcogenides (TMDCs) are direct gap semiconductors similar to GaAs, but they are otherwise very different. Due to the extreme confinement of the excitons, their binding energies are on the order of 500 meV. Despite the confinement, the inhomogeneous linewidth of boron nitride encapsulated monolayers is fairly low compared to the homogeneous linewidth. We measure that the homogeneous and inhomogeneous linewidths are about 0.25 meV and 1 meV respectively. Localized states are measurable in some samples, though the binding energies of the localized states are tens of meV. Encapsulation can also suppress localized states, and we measure samples with no strongly localized states. We also measure nearly radiatively limited linewidths. We measure these samples in a few configurations to enhance the radiative linewidth and thereby determine the radiative linewidth.

CHAPTER IV

Nonlinear Spectroscopy with Thermal Light

4.1 Introduction

This chapter is a somewhat disconnected from the other chapters in the thesis, but this project formed how I understand coherence and motivated the spectroscopic techniques discussed in Chapter II. This work can also be considered a conclusion to a substantial effort by a number of previous group members: Ryan P. Smith, Georgios Roumpos, Andrew E. Almand-Hunter, and Marc Aßmann. This work was motivated by a theoretical prediction that the quantum statistics of light should influence the many-body interactions in semiconductor nanostructures [2, 64, 65]. The moonshot goal was to test if it is possible to directly generate a quantum degenerate exciton distribution, an exciton condensate, by resonantly exciting an exciton population with thermal light.

There are two ways in which the quantum statistics of light have been considered to affect the interaction of light and matter. 1) The most straightforward way of considering the relevance of quantum statistical fluctuations is through the nonlinear optical response of a material. This can be understood by comparing the nonlinear responses of two light sources with an average intensity of I_0. The intensity of the first light source is always I_0, and the other light source fluctuates between 0 and $2I_0$. Of course any response that does not scale linearly with intensity will be different

for these two light sources, as is the case for any nonlinear response. Let's consider second-harmonic generation, which scales quadratically with intensity, as an example. The second-harmonic signal would be proportional to I_0^2 for the first source. The average second harmonic generation from the fluctuating source would be proportional to $\frac{1}{2}[0^2 + (2I_0)^2] = 2I_0^2$. More advanced uses of this application of quantum optical fluctuations are discussed later in this chapter, and much more advanced uses have been used to measure high order correlations in the nonlinear optical response of GaAs QWs to measure the signature of quantum droplets [19, 66]. 2) The less straightforward prediction is that thermal light, a light source with high temporal correlations that is *relatively* straightforward to generate in the lab, could directly excite a population of excitons in a quantum well without creating a polarization. Examples of light sources having thermal quantum statistics are black-body sources and spontaneous emission. A common way of motivating this work has been to note that photoluminescence, the light emitted by the radiative decay of excitons directly from a population state to the ground state, has thermal quantum statistics. Thus by using light having the same quantum statistics the inverse process of directly exciting a population of excitons in low momentum states could be achieved. The predicted effects of thermal excitation, as compared to coherent excitation, are long-range order, anomalous reduction of Coulomb and phonon scattering, and enhanced and directional PL [64].

4.2 Temporal Coherence

Coherence is the correlation between various quantities of a single wave typically calculated as a function of space and time. To introduce the idea of coherence it is important to begin with a discussion of correlation functions, and we will discuss this in the context of temporal coherence.

The first-order correlation function is a measure of amplitude fluctuations. This

function $g^{(1)}(\tau)$ is defined:

$$g^{(1)}(\tau) \equiv \frac{\langle E^*(t)E(t-\tau)\rangle}{\langle E^*(t)E(t)\rangle}, \tag{4.1}$$

where E is the complex field amplitude. The function can be measured using an interferometer in which the field is split into two paths, one path is delayed τ with respect to the other, and the fields are recombined on a photodetector. The function can be thought of as a measure of the visibility of the interference term between the fields passing through both arms. Though general correlation functions can be applied to different arbitrary sources, the first order correlation function in optics specifies that it is between a source and itself and defines the source's degree of first-order coherence. For a single spatial mode (where it is not necessary to worry about the spatial coherence) it is apparent that $|g^{(1)}(\tau = 0)| = 1$ for all fields. For light having a finite bandwidth, the autocorrelation of that light as measured by a photodetector can be written:

$$I_{PD}(\tau) = \int_0^\infty I(\omega)(1 - cos(\omega\tau))d\omega, \tag{4.2}$$

where $I(\omega)$ defines the spectral intensity envelope of the light source. The interference term may be identified as the Fourier cosine transform if $I(\omega)$, and so we realize the relationship between the coherence time, τ_c, and optical bandwidth, $\Delta\omega$, which is $\tau_c = \hbar/\Delta\omega$.

Earlier we noted that in order to neglect spatial coherence, we consider the measurement of a single spatial mode. In quantum optics it is simpler and often necessary to consider states that are fully single mode, which also requires isolating a single temporal mode. A single temporal mode is defined for the duration over which $|g^{(1)}(\tau)| \approx 1$. This constraint becomes one of the most difficult for generating quantum statistical light sources.

The second-order correlation function measures intensity fluctuations within a

single mode (coherent to first-order). This is the measure of the statistical properties of light. It is defined for an intensity measured as a function of time:

$$g^{(2)}(\tau) \equiv \frac{\langle I(t)I(t-\tau)\rangle}{\langle I(t)\rangle^2}. \tag{4.3}$$

For a coherent source whose intensity does not vary over time $g^{(2)}(\tau) = 1$ for all τ. For a thermal, or incoherent source, whose intensities fluctuate chaotically, light is bunched in such a way that the intensity is highly correlated at $\tau = 0$. Thus $g^{(2)}(0) = 2$ and $g^{(2)}(|\tau| \gg \tau_c) = 1$. Examples of incoherent sources include incandescent bulbs, sunlight, and spontaneous emission.

The second-order correlation function is measured by splitting a beam into two paths, measuring the two beams on two separate detectors, and delaying the optical path length of one beam with respect to the other. A Hanbury Brown-Twiss interferometer [67], which is commonly used to measure single photon sources (characterized as having a $g^{(2)}(0) = 0$) having long coherence times on the order of ns, correlates the measurement time of single photons on avalanche photodiode detectors (or photomultiplier tubes) in each arm [68, 69]. The timing resolution here is determined by the speed of the electronics, which is typically greater than 40 ps. Thermal light emitted from narrow resonances of a mercury gas lamp has also been measured with these types of detectors [70], but these too have very long coherence times. The thermal sources we measure here have coherence times that are typically less than 1 ps, and so we require a significantly higher timing resolution than can be offered by a classic intensity interferometer. We instead use balanced homodyne detection, which has been described in detail by G. Roumpos and S. T. Cundiff [71]. The balanced homodyne detection we have implemented uses an ultrafast LO pulse to measure the temporal correlation of noise with a timing resolution given by the duration of the LO [72, 71]. We can easily relate the noise measurement correlation to $g^{(2)}$ by rewriting

Equation 4.3 for $I(t) = \langle I(t) \rangle + \Delta I(t)$:

$$g^{(2)}(\tau) = 1 + \frac{\langle \Delta I(t) \Delta I(t-\tau) \rangle}{\langle I(t) \rangle^2}. \tag{4.4}$$

4.3 Spatial Coherence

Correlation functions are defined with respect to space as well. Of course emission from a point is exactly correlated with itself, but light increases the range of coherence. As the coherence time is related to light source's spectral bandwidth, the spatial coherence is related to the light source's momentum distribution. Since the momentum distribution is limited by the momentum of a photon, the emission from two points separated by less than half the light source's wavelength are correlated. A single mode light source having a very narrow momentum distribution Δk can then have a large spatial resolution given by $\Delta x = 1/2\Delta k$.

One related concept that is important to understand for working with incoherent sources is etendue. Etendue characterizes how light is distributed in angle and area. The quantity, which is the product of the source area and solid angle into which light is emitted can never increase in an optical system where optical power is conserved. This means that the intensity of emission at the source of an incoherent light-emitting object cannot by exceeded by any passive optical system. One can also understand from etendue that in order to maximize the excitation at the sample it is essential to collect light from the emitter with as high a NA as possible, and it is important to image that light onto the sample with as a high a NA as well.

4.4 Creating a Population without a Polarization

Numerical calculations predict a stark difference between the quasiparticle excitations of semiconductors resulting from coherent and thermal excitations [64, 2]. In Fig. 4.1a and b the predicted optically active exciton population is plotted as a solid

line for coherent and thermal excitations, respectively. The delayed excitation and significantly decreased optically active excitation density of the coherently excited system are attributed to the polarization-to-population conversion. In the coherently excited system, the excitation creates a polarization, and the exciton population is created by the interaction of that polarization with acoustic phonons. The population is then widely distributed in momentum, shown in Fig. 4.1c, and the lowest momentum states are more quickly depleted by radiative recombination. For the pulse having thermal quantum statistics, a population is created without the creation of any polarization. This population is directly injected into the low momentum optically active states since the excitons acquire the momenta of the photons. These states still rapidly decay by radiative recombination, but the optically active states still dominate the excitation. In Fig. 4.1d the exciton-exciton pair-correlation function is plotted showing how the narrow spread in momentum leads to much longer range interactions in the sample excited by thermal light.

We wanted to experimentally test the validity of these numerical calculations, but first we had to relate it to a physically viable experiment. On paper it is possible to create a pulse envelope in the time domain and describe that pulse's quantum statistics with an independent and arbitrary function. Experimentally, this is not so easy. A pulse is generated by the constructive interference of a broad range of frequencies. A thermal source, however, has a random and uncorrelated spectral phase. Though it is seemingly the randomness of this spectral phase that prevents the formation of a polarization in the semiconductor, this random phase also prevents the experimental realization of a pulsed thermal light source. We thus cannot exactly reproduce the conditions of the simulation, but we have tried and here present the closest experimental reproduction to date.

For the part of this story that is not mine to tell, see [18] Chapters 6-8. R. P. Smith spent a portion of his PhD using a pulsed laser and pulse shaping to mimic

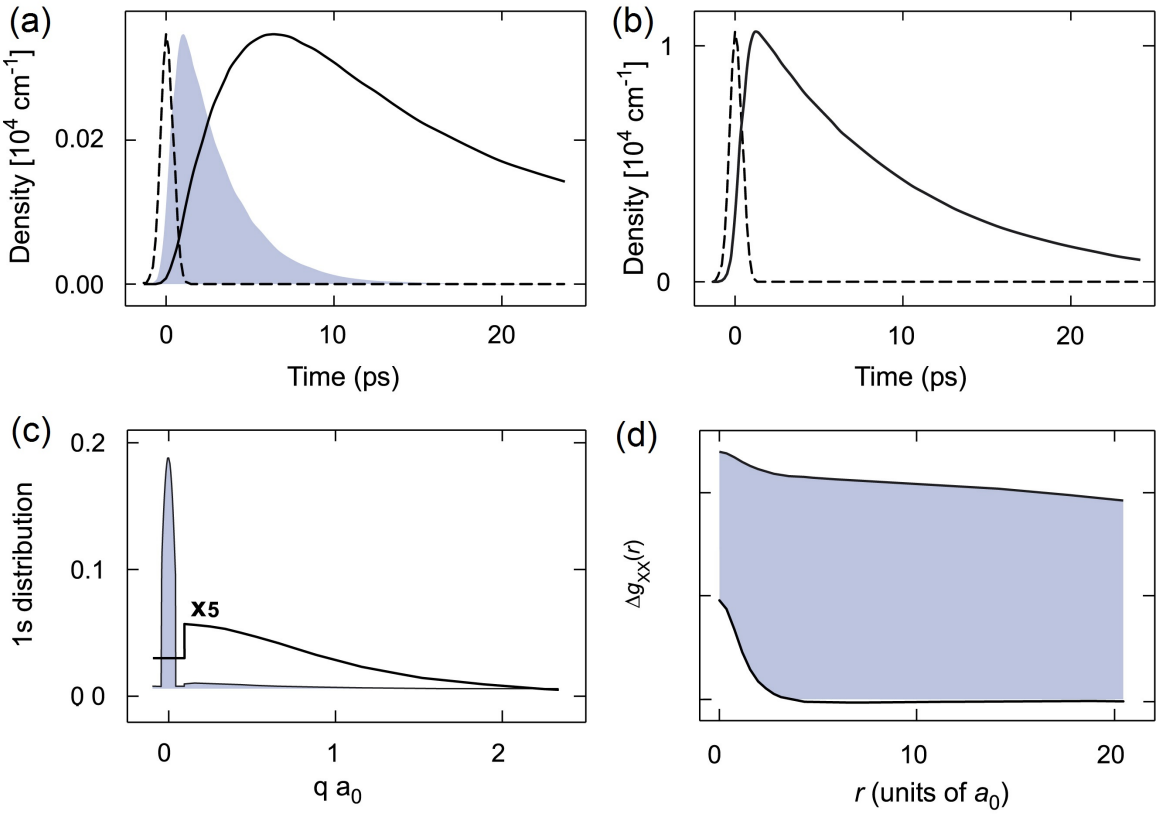


Figure 4.1: Induced polarization (shaded) and optically active exciton population generated (solid line) by pulse excitation (dashed line) having coherent (a) and thermal (b) quantum statistics. (c) Distribution of exciton population for coherently excitation (solid line) and thermal excitation (shaded). (d) Exciton-exciton pair-correlation function showing enhanced long range interaction for thermal excitation (shaded) versus coherent excitation (solid line). Figure from [2].

shots of a single-mode thermal light source. In the following sections we instead start with continuous bright nearly thermal and thermal light sources and compare their nonlinear responses to that of a continuous coherent source.

4.5 Measuring the Bright Exciton Population with Differential Absorption

As discussed in Chapter II, we use heterodyne detection to isolate a nonlinear signal from the resonant excitation beams that are collinear with the signal and also impinging on the photodetector. For the following experiments the optical pump is

actually a continuous or quasi-continuous beam that excites an integrated population in the GaAs sample. Since the population decay time (measured with coherent and pulsed excitation) is approximately 100 ps, the exciton population at any time has been excited by the integrated light absorbed in the 100 ps before that time. We probe the exciton population with a coherent pulse and measure the differential absorption of the probe pulse with the pump on and off. By amplitude modulating the pump beam with a high speed optical chopper, we measure the differential nonlinear signal with a high signal-to-noise ratio (SNR). We heterodyne the nonlinear signal with a LO and scan the LO over the signal in time with phase sensitivity, as described in Chapter II. We Fourier transform this signal to acquire a spectrum of the nonlinear signal. The spectrum is sensitive to many-body interactions [8]. We expect a many-body interaction signal to be most sensitive to the differences in the range and phase of exciton-exciton interactions. There are two aspects of the theoretical prediction that we expect to affect the many-body configuration. First, no significant polarization should be created in the excitation with thermal light. The polarization introduces many-body interaction terms that theory predicts would not be present for thermal excitation, and therefore would not affect the differential absorption spectrum. Second, the population is all created at the low-momentum states. If low-momentum phase-space becomes filled, we expect stronger long-range interactions since low-momentum states are less localized. This stronger interaction would lead to increased dephasing or shifting of the transient exciton absorption peak. Though these two effects may be somewhat compensating, spectrally resolving the transient-absorption signal maximally sensitive to the excitation created by the pump. We therefore use spectrally-resolved differential-absorption spectroscopy to test whether excitation by a thermal light source does in fact create excitons with enhanced long-range interactions with creating a polarization.

4.6 Attempt at Using a Laser Diode below Threshold as a Thermal Source

We began by attempting to excite a nonlinear response with amplified stimulated emission (ASE) from an ECDL pumped below threshold. The ECDL is a diode laser with an anti-reflection coating on the front surface and feedback provided by an optical grating mounted approximately 1 cm from the diode in the Littrow configuration. The external cavity allows a large range of wavelength tuning and significantly narrows the laser linewidth when the grating is aligned to feedback into the lasing mode.

It had been demonstrated by Georgios Roumpos that triggered pumping of a laser diode led to high shot-to-shot intensity fluctuations of the output leading to $g^{(2)}$ measurements of 2 even for relatively high peak powers [71]. Though we used this source in a pump-probe type measurement to measure a nonlinear response of GaAs, we could not trust these measurements. First, we were unable to generate a comparable pulsed coherent source with a $g^{(2)}$ of 1. Though we could generate a pulse output with a narrow lasing spectrum, the shot-to-shot fluctuations of the amplifier inhibited our measurement of a light source with the statistics of a laser. Second, it was difficult to know that the intensity fluctuations of the "thermal" source could be attributed to quantum fluctuations and not classical noise introduced by the pulsed amplifier. For instance the optical spectra indicated simultaneous excitation of many spectral modes that could have resulted from sweeping a lasing spectrum across some range of the bandwidth.

We next used a direct current excitation of the external cavity laser diode to generate the excitation source. G. Roumpos had demonstrated by measurement of $g^{(2)}$ that when pumped above threshold, the emission was indeed coherent. Near threshold he also managed to demonstrate a noisier state having a $g^{(2)} = 1.7$ [71]. Unfortunately this noisier state was very weak and unable to excite a nonlinear response. However, we did manage to significantly misalign the grating in such a way as to not provide

sufficient feedback to the laser diode for it to lase at the pump thresholds G. Roumpos measured. With this configuration we measured $g^{(2)} = 1.9$ and had sufficient power to excite some nonlinearity in GaAs.

In order to compare the ASE result with coherent lasing, we needed some way of comparing the spectra and ensuring identical spatial modes. To excite the sample with similar spectra that overlapped the exciton we needed a narrow-bandwidth (< 1 nm) spectral filter. Using a spectral filter based on a grating would lead to significant loss of light. A filter based on prisms would not have sufficient resolution since prisms are typically not sufficiently dispersive. We thus built a homemade tunable Fabry-Perot filter that transmitted all of the ASE resonant with the exciton. Shown in Fig. 4.2, the filter was made with two partial reflectors each having a reflectivity of 70%. This provided a finesse of 8.8, which corresponds to the ratio of the cavity linewidth and the free spectral range. The spacing between the reflector was set by an 80 micron thick plastic ring so that the free spectral range of the cavity was approximately 2 THz, or 8 meV. All of this was mounted in a 1 inch lens tube with a retaining ring designed with three set screws pressing on the front partial reflector to control the spacing between the mirrors. This could be used to tune the filter wavelength and ensure the surfaces were parallel.

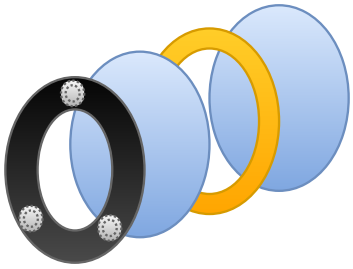


Figure 4.2: Fabry-Perot filter made with two partial reflectors each having a reflectivity of 70% and an 80 micron thick plastic spacer all mounted in a 1 inch lens tube. A large retaining ring is designed with three set screws to fine tune the mirror spacing of the partial reflectors.

We ensured that the sources had similar spatial modes by spatial mode filtering

the beams. A spatial mode filter is made with a short focal length lens and pinhole at the focus having a diameter equal to the diffraction limited spot size calculated for the incoming beam size and lens focal length. Here we used an objective having an effective focal length of 1 cm and a 20 μm pinhole. For a 1 mm input beam diameter, the airy disc diameter at the focus is 20 μm, and so this would be an appropriate pinhole size.

After all this we were able to measure a nonlinear response by pump-probe for both light sources, but the results remained inconclusive. It was fairly apparent that the quasi-thermal light was not exciting a significantly greater exciton population density, but we could not definitively say the responses were identical either.

4.7 Excitation with a Superluminescent Diode

Motivated by ambiguous results and the enhancement of $g^{(2)}$ in a laser diode by an intentionally misaligned external cavity, we purchased a new thermal source. The new light source, produced by Superlum, is a superluminescent diode emitting almost 20 mW of ASE with 20 nm bandwidth out of a single-mode fiber. The spectrum emitted from the diode is a smooth Gaussian with no spectral spikes that would indicate partial lasing. These diodes are made by anti-reflection coating the front facet of the diode and cutting the back facet at an angle so as to prevent any feedback to the emission region. The fiber coupled output replaced the spatial mode filter. Also, we actually had an excess of optical power and were able to to filter the spectrum with a grating based spectral filter, which is more stable than the homemade Fabry-Perot filter.

An ECDL is optimized to have maximum coherence and also coupled into a single-mode fiber so that exchanging light sources is easy, and we can ensure that the single-spatial modes of the two sources are identical. In Fig. 4.3 we plot both the thermal and coherent spectra overlaying the probe transmission spectrum. The transmitted

probe spectrum is spectrally filtered above 1549 meV to not excite the light-hole exciton or heavy-hole continuum states. The absorption dip in the spectrum is the result of absorption by the heavy-hole exciton at 1547 meV and weak absorption by the biexciton is measured at 1545 meV. The coherent and thermal light sources are tuned to the heavy-hole 1s exciton resonance wavelength at low excitation density and spectrally filtered to match the exciton. The coherent source of course has a much narrower linewidth than the exciton, and its measured linewidth here is actually spectrometer resolution limited. Still, both excitation sources are narrow enough to assume they are nearly equally absorbed by the exciton. As the exciton blueshifts with increased excitation density the light sources will no longer be resonant with the exciton. Their spectra are similar enough that we expect comparable absorption both light sources at high densities, excepting effects resulting from their quantum statistics.

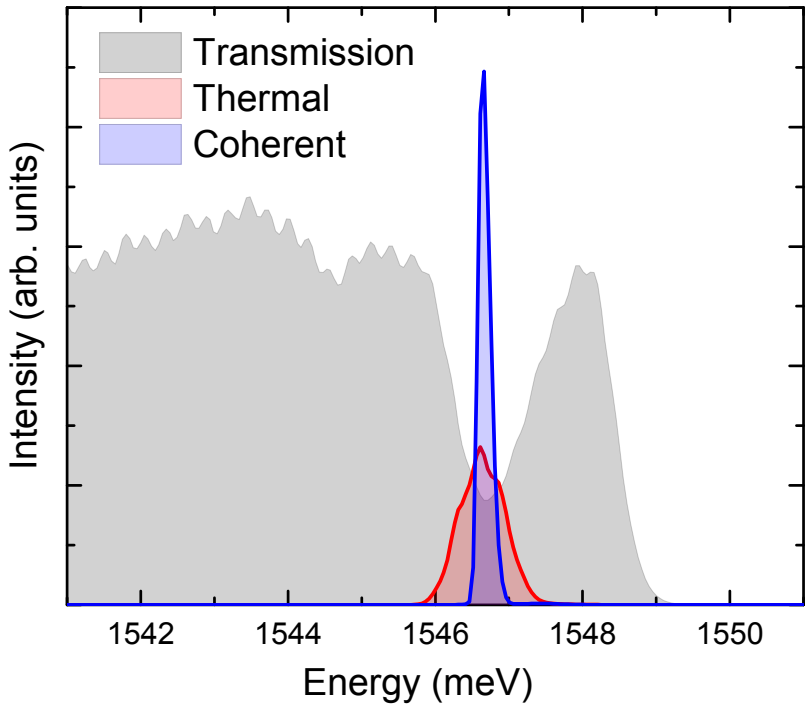


Figure 4.3: Gray shaded region is the probe transmission spectrum. The absorption dip at 1547 meV corresponds to the 1s exciton of GaAs. The shaded red region is the spectrum of the thermal excitation source and the shaded blue region in the spectrum of a coherent excitation source having the same total power.

4.8 Results

We compare the nonlinear signals resulting from excitation by light having coherent and thermal statistics and a pulsed probe sampling the excitation. We plot the real part of the signal, as opposed to the absolute value, because the signal phase is a strong indicator of the types of many-body interactions occurring in the sample. In measurements of the real part of a nonlinear signal, positive values correspond to excited state absorption, and negative value correspond to excited state emission. Absorptive lineshapes, characterized by negative valued wings, are an indicator of excitation induced dephasing and Pauli blocking. Dispersive lineshapes, resembling the derivative of a Lorentzian, are an indicator of excitation induced shift [73].

In Fig. 4.4 we compare nonlinear spectra excited with 80 μW pump beams having coherent and thermal statistics. Since both light sources are coupled into the same single mode, they are focused to identical 700 nm spots. We show that there is no significant difference in the many-body interactions manifest in the two exciton populations. We calculate that the photon density for this measurement in a 100 ps window, corresponding to the population buildup time for the QW exciton, is roughly 6×10^{12} photons/cm^2. We measure nonlinear spectra for three times higher photon densities, which nearly saturates the exciton. We also measure nonlinear spectra at very low excitation densities. We find no differences in the excitation induced shifts or dephasing between the two light sources for the entire range of excitation densities. This is fairly strong evidence that the exciton population distribution generating by these two sources is the same.

The measurement of both low and high temperatures arose from an interest in matching the homogeneous linewidth of the exciton to the thermal response. At 7 K, the homogeneous linewidth is approximately 0.1 meV, measured with multidimensional coherent spectroscopy [41]. At 65 K the linewidth is around 0.7 meV, which is broader than both the thermal and coherent linewidths. Though we see that the

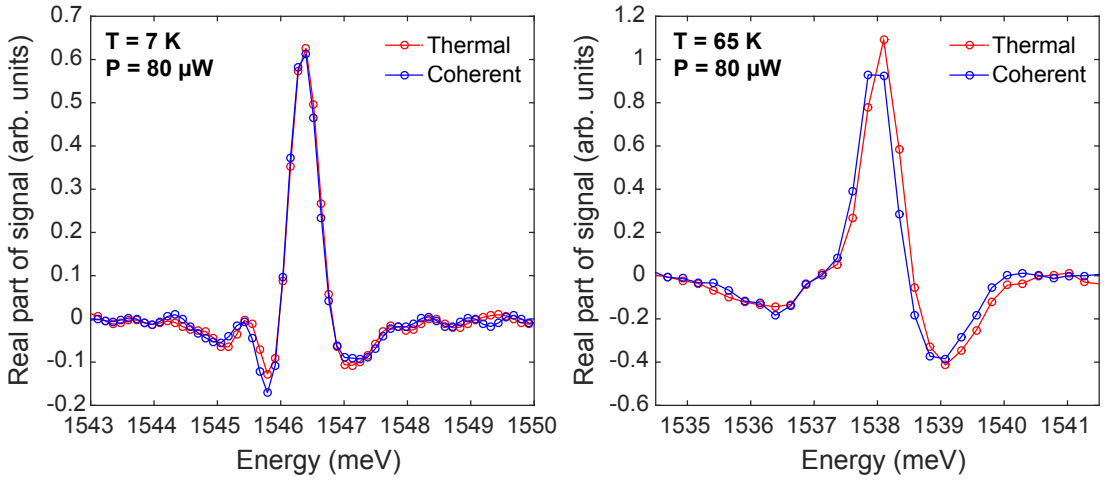


Figure 4.4: Real part of spectrally resolved nonlinear signal. Here we compare nonlinear signals resulting from thermal (red) and coherent (blue) excitation. We show minimal difference in the induced signal for both low (left) and high (right) temperatures.

linewidths are broader at 65 K, we do not see any significant different between the light sources.

The likeness of the differential absorption excited by thermal and coherent light sources is almost surprising. Besides the technical changes of ensuring the beam are spectrally, temporally, and spatially effectively identical, there is a nontrivial explanation for why the statistical fluctuations of the thermal source do not affect the nonlinear response. For excitation by a single mode of each light source there is the difference discussed earlier wherein a light source with high fluctuations samples higher orders of nonlinearity. By using a low noise source to measure the full nonlinear response all the way to saturating the response, we can actually project that measurement onto a state with quantum fluctuations [19, 66]. Projecting the coherent response, R_{coh}, onto that of a thermal state, R_{th}, we find

$$R_{\mathrm{th}}(I_{\mathrm{th}}) = \frac{\int_0^\infty dI_{\mathrm{coh}} e^{-I_{\mathrm{coh}}/I_{\mathrm{th}}} R_{\mathrm{coh}} I_{\mathrm{coh}}}{\int_0^\infty dI_{\mathrm{coh}} e^{-I_{\mathrm{coh}}/I_{\mathrm{th}}}}, \tag{4.5}$$

where I_{coh} is the measured intensity of the coherent light used for the projection and I_{th} is the intensity of the thermal light for which the response is calculated. To accurately determine the projected quantities, we measure the nonlinear response up to very high excitation densities with the coherent light source, plotted in Fig. 4.5. In order to minimize noise on the projection we fit this with a curve that depends linearly on power and saturates, and we use the fitted curve as I_{coh} in Eqn. 4.5. At low powers this nonlinear signal should be dominated by the third order response, which scales linearly with pump power here. So we fit the coherent excitation in Fig. 4.5 with the function $f(P) = \frac{aP}{1+aP/b}$, where a is proportional to $\chi^{(3)}$ and b is the saturation amplitude.

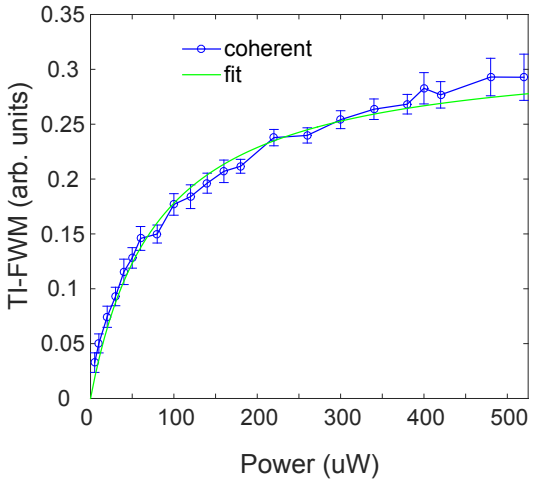


Figure 4.5: Fit of nonlinear response with coherent excitation up to very high excitation. Saturating the response and fitting the response both serve to minimizes error in projecting the coherent response onto a thermal one.

In Fig. 4.6 we compare the time integrated nonlinear signals resulting from thermal and coherent excitation. We also compare these to the projected thermal response for a single-mode thermal state. We find that through careful experimental practice, the response of GaAs to these two sources is nearly identical for low and high temperatures. The deviation of the thermal response from the projected response results from sample integration over many temporal modes of the thermal light. The coher-

ence time of the thermal source is around 1 ps, due to the bandwidth of the spectral filter. This is the time scale on which high intensity fluctuations occur and $g^{(2)} = 2$. The nonlinear signal results from exciton population created over ∼100 ps, which corresponds to integrating over 100 temporal modes of the thermal light source. The exciton population thus averages the high intensity fluctuations of the thermal source so that the nonlinear response is that of a light source with low intensity fluctuations.

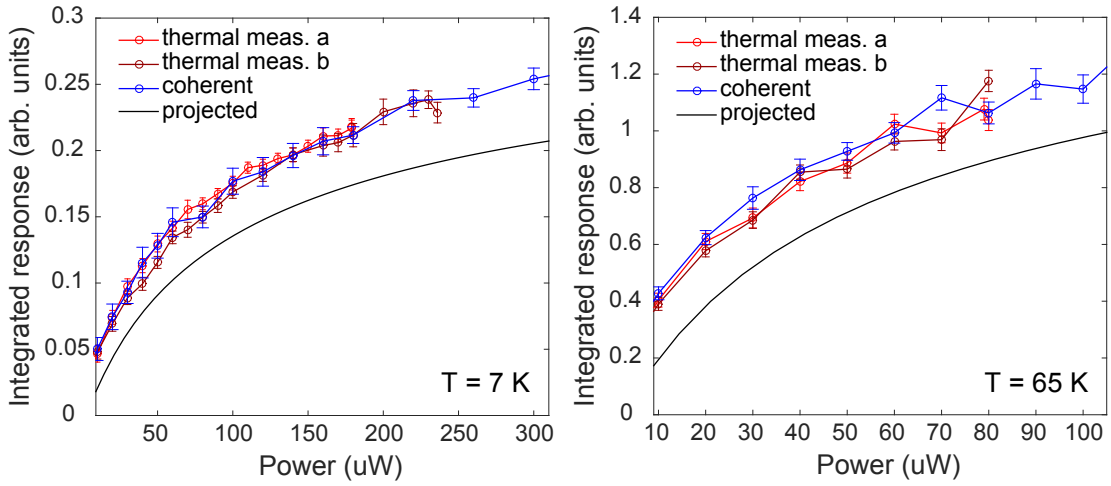


Figure 4.6: Time-integrated nonlinear response as a function of pump power for a thermal excitation source (red and maroon represent two separate measurements) and a coherent excitation source (blue). These are also compared to a projected thermal response (black line) calculated using the measured coherent response. These measurements are shown for low temperature (left) and high temperature (right).

4.9 Conclusions and Outlook

We have demonstrated a nonlinear response induced by thermal light emission collected from a superluminescent semiconductor diode. At the beginning of this chapter we listed two reasons a nonlinear response to thermal light could differ from coherent light, and we measured no difference. 1) The straightforward response, that a source with high intensity fluctuations should sample a wider range of nonlinear response for the same average power constant source, was not adequately tested. Demonstrating

this with a true quantum statistical source requires that the nonlinear response be induced by a single mode of the source. Since we excited a nonlinear response with multiple temporal modes, we could not test this hypothesis. However, its effect can be known by projection, which has been used here and elsewhere [19]. 2) The prediction that thermal light should prevent polarization-to-population conversion has been tested and we have demonstrated that this predicted effect is not significant for light having a coherence time greater than the exciton. There are a few distinctions between the numerical calculations and the experimental test, but these distinctions result from components of the theory that are very difficult to realize experimentally.

Here we list a few possible explanations of the discrepancy that may be applied to the existing theory to find agreement. First, thermal light filtered by the exciton is still coherent to first order for the macroscopic polarization timescale, corresponding to roughly 1 ps. So some first order polarization is still created on short time scales, though the polarization is not built up over long time scales. Second, impulsive nonlinear excitation with a thermal light source is not currently feasible experimentally, but this could lead to significantly different results. Third, it is possible that exciton localization leads to low spatial coherence and dominates the effect of the enhanced long-range exciton interaction predicted by theory. However, this seems unlikely since localization length scales are typically on the order of 100 nm, which is quite large compared to the exciton Bohr radius and the values plotted in Fig. 4.1d. Lastly, it is possible that averaging over coherence times degrades this effect as well, but this too seems unrelated to the proposed theoretical explanation of the effect.

Moving forward, we see a couple more tests to do. First, it would be interesting to repeat this experiment with thermal light having a short coherence time. The difficulty is that comparison to a coherent source would then require the coherent source have a broad bandwidth as well, which would lead to the comparison difficulties R. P. Smith encountered in his studies [18]. The closest way to make them comparable

would be to compare the broadband thermal response to that of a very high repetition rate coherent laser. Microresonator-based lasers have been demonstrated with repetition rates between 10 and 1000 GHz [74]. Any repetition rate in this range would correspond to multiple pulses within the integration time of the exciton population. Therefore lasers with such high repetition rates would be strong candidates for a comparable coherent source.

As another test, it would be interesting to measure a nonlinear response with a single mode of thermal light that differs from that of coherent light. Since the power of a superluminescent diode is very high, it should be possible to measure such an effect in a system with a response time that is on the order of the thermal coherence time or faster. One test we propose would be to measure integrated second harmonic generation of the two pump sources. These could either be focused tightly into a nonlinear crystal or mixed with a coherent pump pulse. In either case one would expect an enhanced nonlinear response from the thermal source as compared to the coherent source due to the high intensity fluctuations. A second test would be to do the same pump-probe experiment presented here on monolayer molybdenum diselenide. As we demonstrate in Chapter VII, the majority of the exciton population in this material decays within 1 ps due the minimal scattering to high-k states. Therefore the response time is on the order of the coherence time of the thermal light we have generated here. The nonlinear response of these samples is also very strong due to the enhanced Coulomb interactions resulting from the decreased Coulomb screened by the monolayer. Since we could excite a substantial population with thermal light in a short window of time, the system is a good candidate for revealing differences in the nonlinear response between coherent and thermal light.

CHAPTER V

Optical Tweezing of Excitons

5.1 Introduction

Optical trapping is a powerful tool with applications ranging from cell manipulation to atomic Bose-Einstein condensation [75, 76, 77]. A class of these traps, generally known as optical tweezers [78], exploits the optical density gradient of a single focused laser beam to exert a force on angstrom- to micron-sized dielectric particles in a medium with a low polarizability [79]. If instead the medium were a macroscopic dielectric, trapping particles would require that the particles be more polarizable than the medium. Application of optical tweezing to quasiparticles in a solid medium thus requires that the quasiparticles' polarizabilities be relatively larger than the solid in which the quasiparticles exist. We will focus on quasiparticle excitons, a bound pair of a conduction-band electron and a valence-band hole (electron vacancy). Here we report measurement of a possible optical gradient force on excitons in semiconductor quantum wells that is greatly enhanced by many-body effects. Specifically, we measure the ultrafast spatial evolution of a small (0.7 micron) excitation spot. For a few picoseconds after resonant excitation, we find the excition spot decreases before increasing by diffusion. We present several possible models for this phenomenon as we work with theorists to determine the most likely source of the effect. This work helps identify the most important terms for dealing with inhomogeneous optical fields

in semiconductors, expands the field of optical trapping, and introduces the prospect of trapping direct excitons in quantum wells.

Many quasiparticle traps have been created in semiconductors, but thus far none have been developed that is a corollary to the optical gradient traps typically applied to atoms. Since the first electron-hole liquid was confined using a strain-induced trap [80], exciton confinement has been implemented using strain-induced traps [81, 82], magnetic traps [83], and electrostatic traps [84, 85]. Laser-induced traps have also been implemented [86], though the trap potential was created by exciton repulsion.

The use of the optical gradient that exists at a tight laser focus to trap and manipulate excitons has been proposed [87, 88, 89], but such a force has yet to be experimentally verified. The proposal of transversal light forces in semiconductors included the revelation that a far-off resonance trap in a semiconductor will actually repel an electron-hole plasma [87]. This means that a small force would look similar to diffusion or exciton repulsion. Our theory is consistent with this repulsive force for detuned light, though we have experimentally measured an attractive force on excitons for light that is resonant with the excitonic resonance of the medium. We describe this attractive force using many-body effects.

5.2 Experimental Methods

We have developed an experiment to measure the spatiotemporal dynamics of an exciton population in a sample containing 10 uncoupled gallium arsenide (GaAs) quantum wells (QWs) excited by a spatially small ($\sim$0.7 μm) resonant pump pulse. We achieve such a spot size using a 100x long-working distance microscope objective with a numerical aperture (NA) of 0.5 to focus collinear pump and probe beams onto the QW sample. To measure the size of the excitation spot, we scan the excitation spot across the probe spot by varying the angle at which the pump beam enters the objective, shown in Fig. 5.1. The probe pulse responds to the magnitude of the pump

excitation as a function of the relative spatial positions of the pump and probe. To prevent clipping the pump at the entrance of the objective while scanning the entrance angle, a steering mirror is imaged on the objective. The steering mirror is mounted on a pivot edge with a piezoelectric actuator glued to one side of the mirror. The design uses a piezoelectric actuator with a scan range of 20 μm and provides an angular range of about 3 mrad, which is enough to spatially scan the pump vertically over a range that is about ten times larger than the probe spot. By imaging the focus of the transmitted beams on a CCD with a magnification of 250, we can accurately determine the relative positions of the pump and probe beams at the focus. Temporal resolution is obtained by scanning the delay of the ultrafast pump pulse with respect to the probe pulse using a translation stage in the pump beam's path.

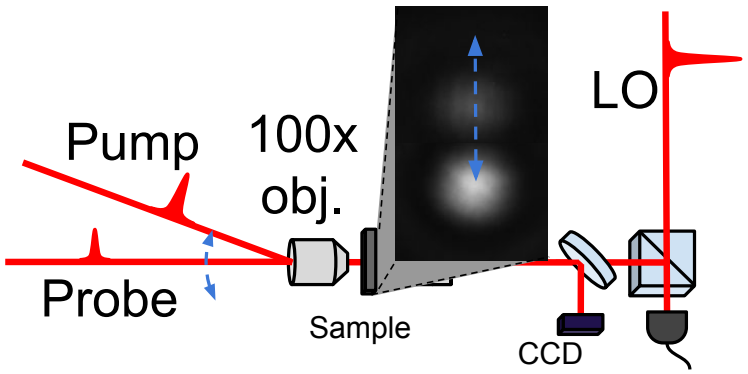


Figure 5.1: A simplified schematic of the spatial scanning portion of the experimental setup. A steering mirror in the pump path is imaged on the entrance of the first microscope objective to scan the angle of entry of the pump beam without changing the position the beam enters the back of the objective. This serves to scan the position of the pump beam's focus without clipping the beam or changing it's k-vector at the focus. The pump and probe beams are focused to ~0.7 μm spots using a microscope objectives with a NA of 0.5. Relative beam positions are recorded for each measurement by imaging the spots on a CCD camera after the sample. The transient absorption signal is measured by heterodyne detection using a LO pulse as described in Chapter II.

Co-circularly and co-linearly polarized pump and probe beams have been compared with identical results between the two. All data plotted here are measured

with linearly polarized light to maintain uniformity. The spectra of both the pump and probe pulses are centered on the heavy-hole exciton and spectrally shaped to not excite the light-hole exciton. After shaping, the spectra have a 16 meV bandwidth that correspond to a bandwidth-limited 120 fs duration.

We measure the probe by interfering it with an intense local oscillator (LO) pulse that does not pass through the sample. The interference signal is isolated from the DC signal by indepently shifting the LO and probe beam frequencies using acousto-optic modulators (AOMs), as described in Chapter II. The LO is frequency shifted by +80 MHz and the probe is shifted with +80.09 MHz, so a beat note between the two at 90 kHz is measured when the two beams interfere temporally on a photodiode. In this way we use the LO to temporally gate the measurement of the transmitted probe as shown in Figure 5.2.

The purpose of the pump beam is to create an excitation spot whose population will affect the sample's absorption of the probe. We heterodyne detect a transient absorption signal almost exactly as described in Chapter II, except we amplitude modulate the pump beam with a sinusoid instead of a square wave. Sinusoidal modulation is sufficient for exciting nonlinear signals up to third order. Though it is not essential here, care should be taken for quantitatively describing signals with higher orders of nonlinearity [8]. To generate sinusoidal amplitude modulation, we use a single AOM driven with two frequencies: -76.92 MHz and -77.08 MHz, where we use the -1st diffraction order to isolate negatively shifted frequencies. We have chosen these frequencies to minimize any unintended interference and noise between signals. The two frequencies beat to give a sinusoidal amplitude modulation of the pump at 160 kHz. The nonlinear signal, which we will call the interaction signal, is measured using a lock-in amplifier referenced at the difference frequency between the AOM-based amplitude modulation and the probe-LO beat note (160 kHz - 90 kHz = 70 kHz).

Using a pump that is significantly more intense than the probe leads to a dominance of local-field effects in the interaction signal that confuses the spatial dynamics. This is discussed in more depth in the Local-Field Effects section of this chapter. To minimize local-field effects relative to actual population dynamics we optimize the relative field strengths for measuring in the $\chi^{(3)}$ regime, where the pump field acts twice to create a population and probe field acts once. Therefore the strength of the pump field (at the maximum amplitude) is set to be twice the probe field strength.

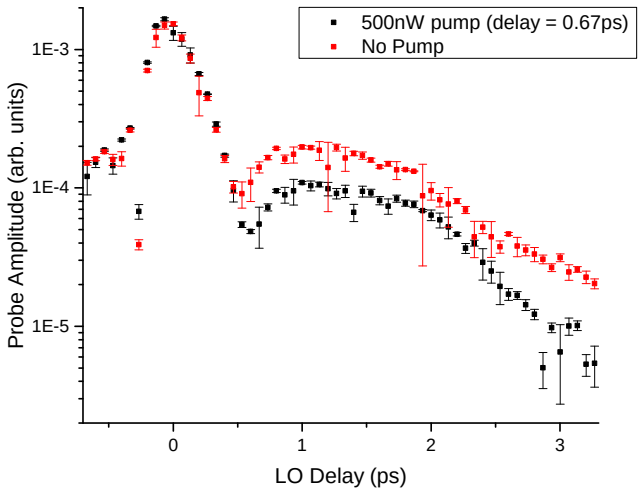
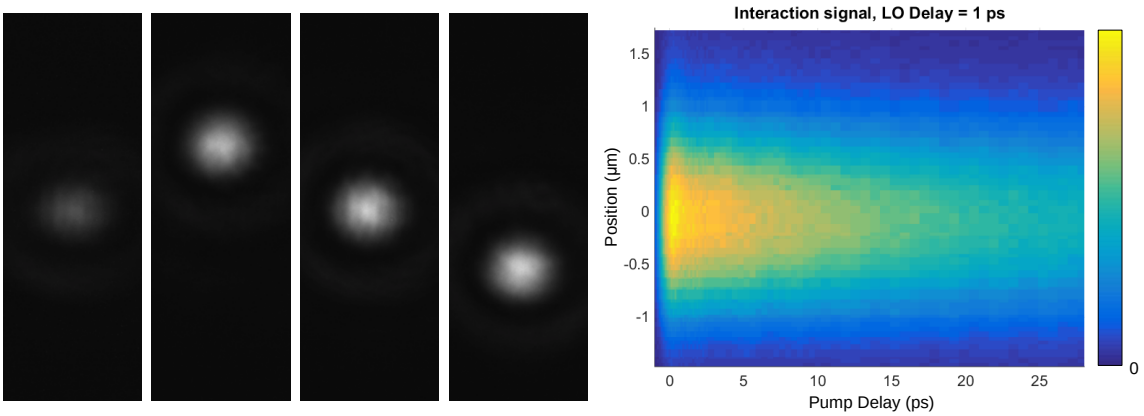


Figure 5.2: The LO is used to make a gated measurement of the probe. When the LO is scanned over the probe, we can temporally distinguish the pulse and emitted coherence. Since the emitted coherence is resonant with the exciton, the greatest nonlinear signal can be measured at the peak of the coherence. The probe-LO delay is therefore typically fixed such that the LO amplifies only this coherence. The figure shows the probe signal when the pump is off (red) and on (black). The interaction signal is proportional to the difference between these signals.

5.3 Analysis

The CCD on which the focus spot is imaged takes an image of the pump and probe at every steering mirror position, as shown in Figure 5.3. We accurately determine each spot position by taking a weighted mean of all pixels above a set threshold. The



(a) (b) (c) (d) (e)

Figure 5.3: Images of pump and probe spots on the sample with a magnification of 250. (a) Image of the probe, repeated for every position but unmoved throughout each scan. (b)-(d) First, middle, and last images of the pump for a scan of the pump position. 75 positions between are used for each determination of the spot size. (e) Interaction signal plotted as a function of position found using images and relative temporal delay between pump and probe.

absolute difference between the pump and probe positions is then used as the position value in the plots below, and the sign of the position shows whether the pump was spatially above or below the probe.

We measure the interaction signal as a function of the pump position found above and of the pump delay. An example plot of this data is shown in Fig. 5.3 (e). We determine the width of the excitation spot at each pump delay by calculating the second moment of the spatial scan at that delay. We calculate the second moment σ using:

$$\sigma(T)^2 = \frac{\int r^2 \, S(r, T) dr}{S(r, T) dr}, \tag{5.1}$$

where r is the measured separation between the probe and pump, and $S(r, T)$ is the interaction signal as a function of the position and relate delay T. For a Gaussian curve, the second moment σ will be equal to the Gaussian width. This procedure has the advantage of not assuming any form of the curve. This is particularly important for highly nonlinear signals where the spatial response to a Gaussian will not be

Gaussian. There are two small artifacts that result from performing a finite integral over the data. First, the width will be underestimated since the full data tails are not considered. Second, this calculation will damp width dynamics. For instance if the width broadens, more signal will be in the tails, and this signal will not be measured. The calculation will still measure an increased second moment, but it will not be as significant. The same is true for spot-size decreases.

5.4 Results

Following excitation by the 120 fs pump pulse, we observe a transient decrease in the size of the excitation spot that persists for 2 to 3 ps. After this time, exciton diffusion dominates, and the spot size increases as expected. We find that the magnitude of the spot-size decrease increases with excitation density, plotted in Fig. 5.4, and that the decrease does not occur for larger initial spot sizes. We measure the signal dependence on excitation pulse energy density from 130 to 980 nJ/cm^2.

The finding that the spot size initially decreases is surprising. It is typically sufficient to calculate local semiconductor responses using homogeneous semiconductor Bloch equations. The center of the distribution is at a higher excitation density than the tails. We would therefore expect that the center of the distribution would decay faster than the tails, which would only cause the excitation spot to grow with time. However, this intuition is not sufficient to convince ourselves that the result is surprising. If the homogeneous semiconductor Bloch equations are sufficient to explain these spatial dynamics, the same experiment with a larger spot size will yield the same results. For this test it is essential to excite the sample with the same optical density, corresponding to increasing the beam power by the square of the spot size increase.

In Fig. 5.5 we pump the sample with a beam that is roughly 2 μm, which is almost three time larger than the 0.7 μm spot used in previous measurements. The beam

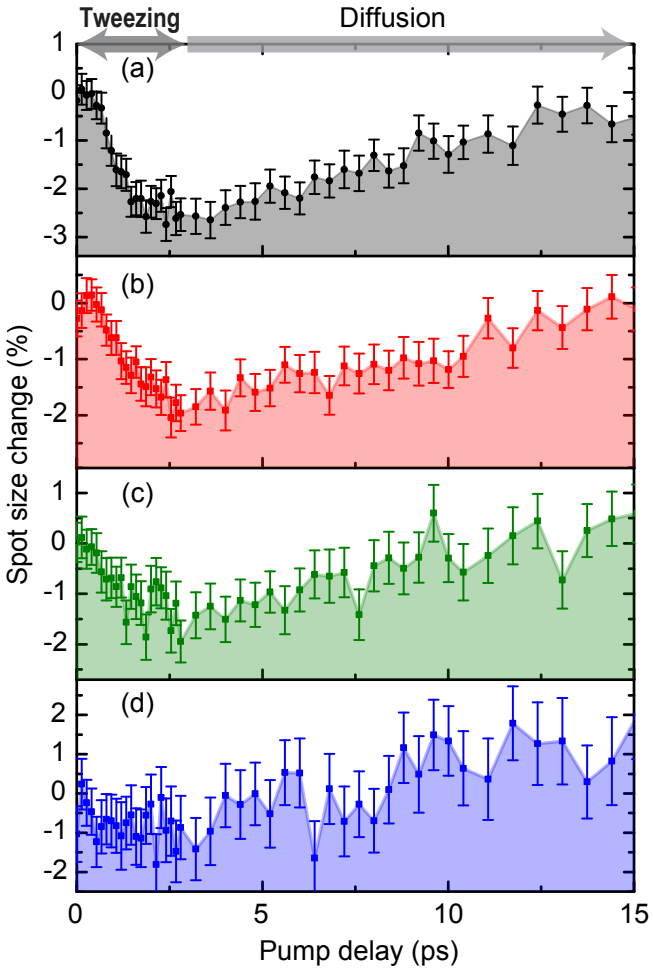


Figure 5.4: Excitation density dependence of tweezing for excitation pulse energies in units of nJ/cm^2: (a) 980, (b) 670, (c) 540, and (d) 270. We find that the magnitude of the spot size decrease increases with excitation. The subsequent spot size growth in each data set is due to the expected diffusion of excitons. The error bars indicate standard deviation for each point.

power corresponds to an excitation pulse energy density of 600 nJ/cm^2. First, we see the linear rate of diffusion is much lower for the large excitation. The linear rate of diffusion is actually nearly a factor of three lower. Since diffusion leads to a constant increase in the spot size area with time, the rate of increase of the spot radius scales inversely with the spot radius. The diffusion effect therefore obviously depends on the spot size, and we are sensitive to that change with this experiment. Within the error bars, we also measure no significant spot-size decrease at low pump delays. Therefore this effect also depends on spot size and cannot be explained just by considering local

responses to the inhomogeneous excitation.

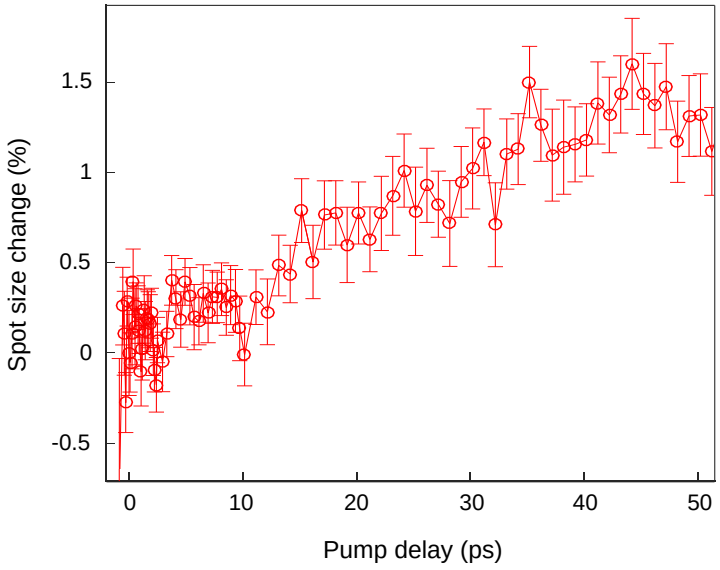


Figure 5.5: In contrast with effects measured using a diffraction-limited spot, a large spot size does not induce any tweezing effects. We use a 20x microscope objective with a NA of 0.4 to create a ∼2 μm spot. Pumping the sample with the same excitation density as we did for the small-spot measurement does not induce a spot size decrease at low times. The linear rate of diffusion also decreases, as expected.

Given its timescale, where the peak of the induced polarization occurs at the same delay time as the peak in the derivative of the spot size, we are inclined to attribute the initial spot-size decrease to a gradient force applied by the radiating macroscopic polarization on the sample excitation. This is counter-intuitive because the polarizability of non-interacting particles on resonance is zero. Classical tweezing is applied to particles with positive polarizabilities relative to the medium in which they exist, and so the tweezing light must be red detuned from the resonance. Tweezing an excitation, however, is not so simple. In the following section we present a few possible explanations for the spot-size dependent effect we measure in resonantly excited GaAs QWs.

5.5 Simplest Theories that Explain the Result

These models have a dependence on the spot being small because either they 1) depend on the gradient of a the field/polarization or 2) result from wave packet evolution dictated by the uncertainty relation.

The first two sections below, the semiclassical model and excitation-induced shift (EIS) model, are both theories that actually correspond to optical tweezing of excitons. Though presented as separate theories, they are really two sides of the same coin. The third model is based on a set of terms in the inhomogeneous semiconductor Bloch equations that our theory collaborators identified as relevant.

5.5.1 Semi-classical model

This model basically relates what kind of exciton dipole would be required to measure the spot size decrease that we measure if it results from the radiated macroscopic polarization optically tweezing a dipole. The model considers a force that is proportional to the induced polarization and includes particle diffusion. It is a very simple model to motivate a microscopic explanation.

Tweezing or trapping forces, originally proposed by [90], can move neutral particles in a light field with a nonuniform transverse profile. The most common realization of the force is at the focus of a Gaussian beam where the electric field gradient forces neutral particles radially inward. The application of the traversal force to trapping and moving particles is called optical tweezing. The regime in which particles are much smaller than the wavelength of light is known as the Rayleigh regime. Here the transversal force is given by $\vec{F} = (\vec{p}\cdot\vec{\nabla})\vec{E}$, a detailed calculation of which is presented in the Calculations section of this chapter. The force scales linearly with the tweezed particle's dipole moment ($\vec{p}$), the inverse square of the spot size, and the square root of the beam power.

We estimate the tweezing field ($\vec{E}$) by measuring the 3rd order polarization, shown

in the inset of Fig. 5.6. We measure the spot size, and use literature values for the effective exciton mass [91]. The fit in Fig. 5.6 uses iterative calculations of the position of a polarizable medium with polarizability α that is accelerated by $A(T) \propto E(T)$:

$$\begin{aligned} r_{i+1} &= r_i + \left(\frac{D}{2r_i} + v_i\right)\Delta T, \\ v_{i+1} &= v_i + \left(\alpha A(T) - \gamma_D v_i\right)\Delta T, \end{aligned} \tag{5.2}$$

where r_0 is the initial spot size from the data, D is the diffusion constant that is well constrained by the data, γ_D is the damping constant accounting for scattering that opposes exciton motion, and ΔT is a time step along the pump delay T. Using this model we get a remarkable fit to the data.

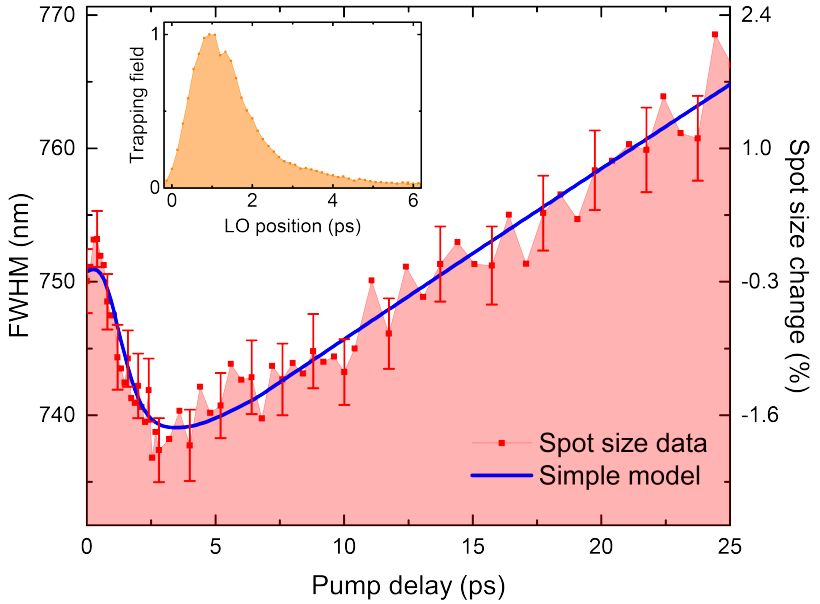


Figure 5.6: Simplified tweezing model relates that the polarization drives the excitation-spot size decrease. We measure the electric field radiated by the polarization, plotted in the inset. Using this field as the source of tweezing (along with a motion damping coefficient to which the model is very robust, and a well constrained diffusivity) we numerically model the spatial dynamics of the excitation for each measured excitation density. Here we plot one of the data sets and fitted model with a damping coefficient of 4 ps^{-1}. Representative error bars indicating standard deviation are shown.

The fit agrees well with the data for all intensities. We find the fits are fairly

robust for a damping constant in the range of 2 to 6 ps^{-1}. This means that the magnitude of the tweezing force applied to a dipole is not very well constrained by the model, but we can estimate that the peak force on an exciton for the 670 nJ/cm^2 would have to be roughly 1 fN to account for the spot-size decrease that we measure. To determine the exciton dipole necessary for this force, we plug in the peak electric field strength that we calculate using the excitation beam power and other measured quantities, described in the Calculation section of this chapter. We find an electric dipole $p \sim 70$ debyes, which corresponds to $p/e \sim 1.5$ nm. This value is greater than the dipole moment of an electron-hole pair without considering Coulomb interactions, but it indicates that measuring a gradient force on excitons would not be entirely unreasonable.

There is also significant precedent to use a Coulomb-enhanced dipole moment to calculate the properties of excitons in semiconductors [50, 92, 93]. Since quasiparticles in crystals are delocalized, excitons are considered collective excitations of many electron-hole pairs. This can significantly enhance the dipole moment of a coherently created exciton, and so we are motivated to consider Coulomb interactions between electron-hole pairs in order to describe a possible tweezing effect. The Coulomb enhancement of the Rabi energy is generally considered to be about a factor of 2 beyond that of a single electron-hole pair [92]. If we consider the enhanced dipole, we must consider effects on the coherently created polarization, and so the orientation of the dipole is determined by the excitation field. Confirmation of the effects described here requires a more rigorous treatment and most likely support from a microscopic theoretical treatment.

5.5.2 Excitation-Induced Shift Model

We propose a physical model, illustrated in Fig. 5.7, that relates how the susceptibility of the tweezed material *must* differ from the background susceptibility.

Here we relate the susceptibility of the excitation (the tweezed material) directly to the nonlinear susceptibility. This is justified by considering that the linear susceptibility changes with sample excitation, so we say the linear susceptibility of the excitation is the difference in total susceptibility with and without excitation: $\chi(\text{excited GaAs QW}) - \chi(\text{GaAs QW}) = \chi(\text{excitation})$. This quantity is really equivalent to the nonlinear susceptibility. Typically the nonlinear response of semiconductors is dominated by many-body effects [7]. Typically phenomenological many-body terms are excitation induced broadening and blue shifting of the exciton resonance, called excitation-induced dephasing (EID) and excitation-induced shift (EIS), respectively. If we determine the excitation susceptibility for a sample whose nonlinear response is dominated by EIS, we find an excitation susceptibility that should be tweezed by the radiated polarization.

In contrast to previously explored gradient forces in semiconductors [87, 88], this model does not involve direct tweezing of electron-hole pairs. It arises from considering the role of many-body effects on the linear susceptibility.

5.5.3 Confined wave function model

One final theory relates the source of the tweezing effect to the uncertainty principle. Though this would typically lead to a spot-size increase, some non-Gaussian distributions actually evolve to have a narrower spot size. We can very directly relate this possible source of spot-size decrease to the evolution of a flat-top beam in optics. As a flat-top beam propagates it will first evolve into a narrower Gaussian beam and then into a wider Gaussian. The same is true for any super-Gaussian. Thus the "tweezing" result could be due to the creation of a super-Gaussian excitation, resulting from saturating the sample at the distribution peak, that evolves into a narrower Gaussian shape.

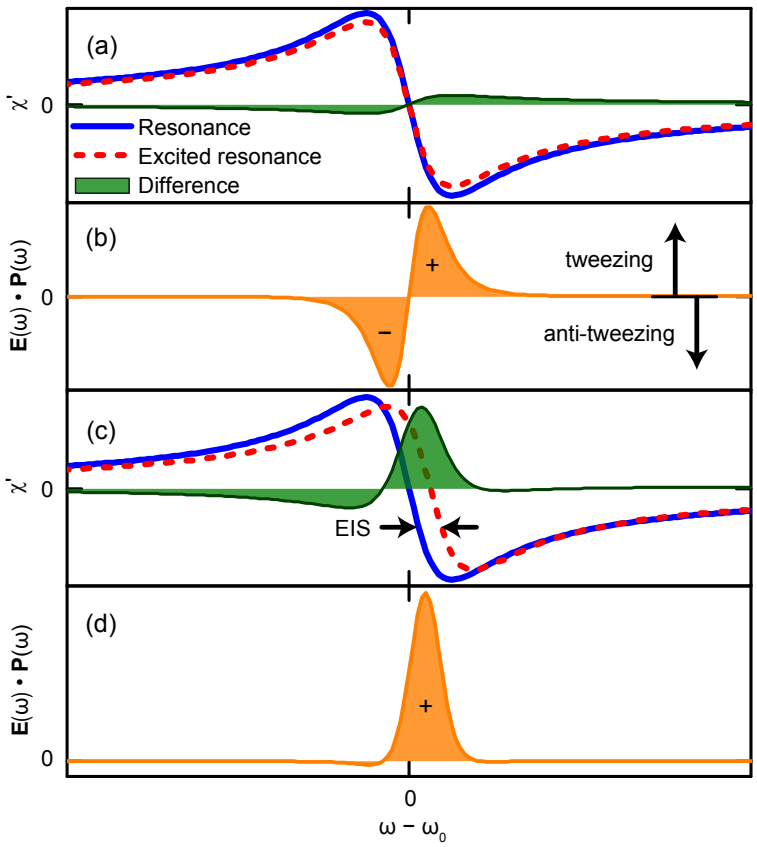


Figure 5.7: The real part of the electric susceptibility (χ') of particles in a medium must be greater than that of the medium for tweezing to occur. In **a**, we plot (solid blue line) the χ' of an unexcited resonance centered at 0. After excitation by a pump pulse the polarizability is saturated to a small degree (dashed red line). The difference (green shaded region) is the χ' of the excitation. In **b** we show the product of the tweezing field (**E**) and the polarization density (**P**) of the excitation calculated using χ' above. This integrates to zero, and so one would not expect tweezing for this resonant case. In **c** we plot the same curves as a, but we also consider the excitation-induced blue-shift of the energy (EIS). In **d** we show that the huge enhancement of the excitation χ' by EIS leads to a large tweezing force that would not be possible without the many-body EIS effect.

5.6 Local-Field Effects

There are a number of considerations that we have made to justify our interpretation of the experimental results. In particular we have explored higher-order nonlinear effects and are careful to minimize their effect on our measurement of the spatiotemporal dynamics of the exciton population.

We would like for a probe measurement such as this to only be sensitive to the population created by the pump, but this is not the case. The collinearity of the pump and probe means that any four-wave mixing (FWM) signal between the pump and probe will go in the same direction as the probe. It important to realize that the amplitude modulated pump beam can be equivalently described by two electric fields that have been frequency shifted by the two driving frequencies. For low temporal delays between the pump and probe, we must therefore consider signals resulting from alternative pulse ordering. In the perturbative description of nonlinear optics [94], the frequency at which we are measuring makes it possible to measure any FWM signal emitted when the pump acts twice (pump and conjugate), and the probe acts once.

In Fig. 5.8 we plot response and spot-size dynamics measured for excitation of the heterodyne detected transient absorption signal using a pump power that is about ten times greater than the probe power. We measure spurious effects on ps timescales that we describe below as resulting from local-field effects. We also demonstrate that these spurious effects depend on the delay between LO and the probe. We therefore have a way to check whether or not the experiment with any given set of parameters is sensitive to these spurious effects, and so we definitely demonstrate that the spatial dynamics described above are not due to this local-field effect.

To explain the effects in Fig. 5.8, we look at the temporal spacing of the signal amplitude peaks and relate those to the probe-LO delay. We find that all the amplitude peaks are spaced by a factor of 1.1 ps. We can see from Figure 5.2 that the coherence emitted by this system after a pulsed excitation peaks at 1 ps. For a very strong pump pulse, we can actually expect the peak of the polarization induced by the pump to be comparably to the peak of the probe pulse. There could therefore be measurable signals resulting from population creation by the pump and probe pulse, and the pump polarization probing that population. The data all support that the

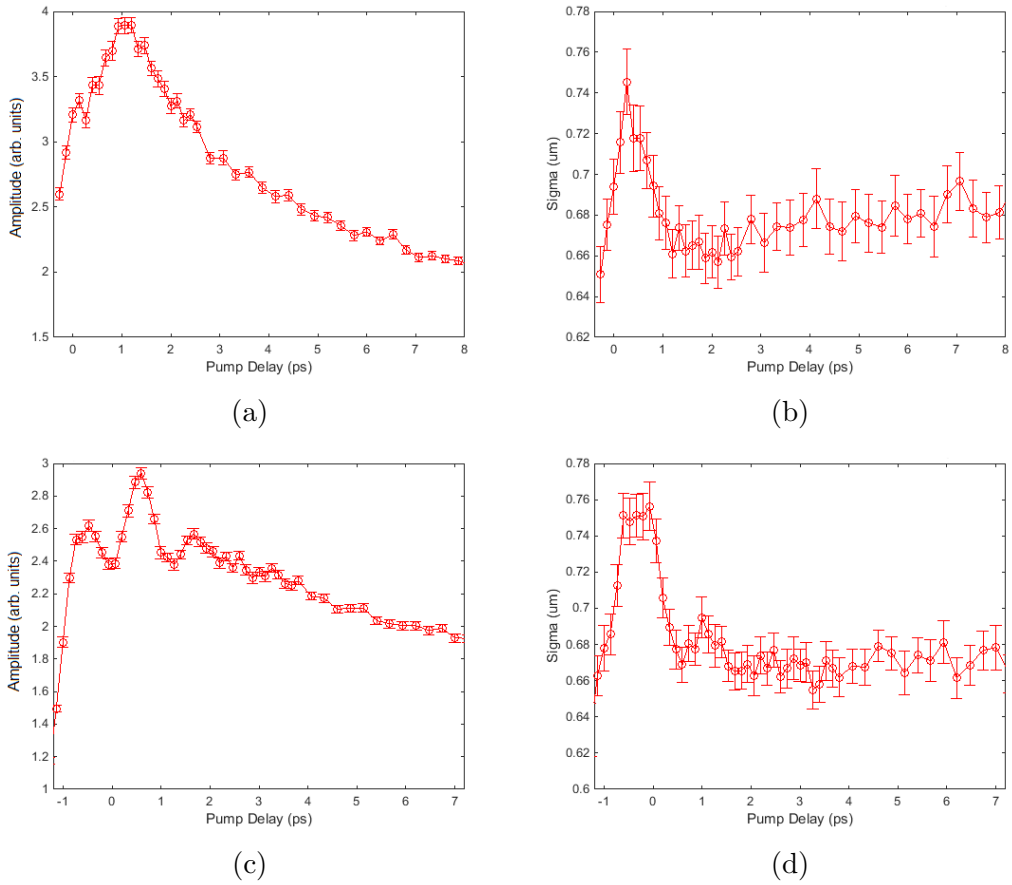


Figure 5.8: Interaction signal as a function of spatial position and pump-probe delay for a pump power $\sim 10\times$ the probe power. Measurement artifacts obscure the signal of interest. (a), (b) Interaction signal amplitude and corresponding width for a probe-LO delay of 1.0 ps. (c), (d) Signal amplitude and corresponding width for a probe-LO delay of 1.9 ps.

sharp peaks are due to the pump coherence acting as a probe, which is one type of local-field effect.

Width dynamics seem strongly correlated to the amplitude fluctuations. In Fig. 5.8 the amplitude peaks almost always correspond to local minima in the width. This can be understood by considering the series of pulses that generate populations giving rise two different signals we seem to measure. 1. The intense pump saturates the exciton population more at the spatial center of the pulse, thereby broadening the excitation spot. 2. The population created by the pump acting only once and probe is not as saturated at the center. The latter, which corresponds to the pump coherence functioning as a probe, therefore has a measurably narrower width.

In order to isolate the desired probe signal, it is necessary to suppress the measurement of local-field effects. This is accomplished by optimizing the field strengths for measurement of a $\chi^{(3)}$ signal, and setting all interacting fields equal. Since the pump acts twice, $E_{\text{pump}} = 2E_{\text{probe}}$. This means that the contribution of the pump coherence to the measured probe signal should be at least an order of magnitude smaller than the desired signal when the LO is fixed at the peak of the probe coherence. Additionally, experiments are all performed with the LO-probe delay set on a few different points to ensure that interesting effects are independent of this delay.

5.7 Calculations

Here we present calculation used for the tweezing model described above. The force on a dipole ($\vec{p}$) in an electric field with a time dependence, a Gaussian profile, and amplitude E_o is given by:

$$
\begin{aligned}
\vec{F} &= (\vec{p}\cdot\vec{\nabla})\vec{E} + \frac{d\vec{p}}{dt}\times B \\
&= \alpha[(\vec{E}\cdot\vec{\nabla})\vec{E} + \frac{d\vec{E}}{dt}\times B] \\
&= \alpha[-\frac{1}{2}\vec{\nabla}E^2 - \vec{E}\times(\vec{\nabla}\times\vec{E}) + \frac{d\vec{E}}{dt}\times B] \\
&= \alpha[-\frac{1}{2}\vec{\nabla}E^2 - \vec{E}\times(-\frac{d\vec{B}}{dt}) + \frac{d\vec{E}}{dt}\times B] \\
&= \alpha[-\frac{1}{2}\vec{\nabla}E^2 + \frac{d}{dt}(\vec{E}\times\vec{B})],
\end{aligned}
\tag{5.3}
$$

where α is the polarizability of an isotropic material that induces the dipole $\vec{p}$.

Here the first term corresponds to tweezing, and the second term correspond to scattering in direction of the light. This latter term normally dominates resonant excitation, but is opposed by the quantum well confinement.

The first term can be applied to a polarizable material, where α is determined from the susceptibility. Alternatively it can be applied to a fixed dipole $\mathbf{p}$, where we can use the fitted curves in Fig. 5.6 to determine the dipole of the tweezed particle. For this second option we break down the first term into mostly known or measurable parts:

$$
\begin{aligned}
\vec{F}_{tweezing} &= (\vec{p}\cdot\vec{\nabla})\vec{E} \\
&= \frac{E(r)}{r}[\vec{p} - \hat{r}(\vec{p}\cdot\hat{r})] + \hat{r}(\vec{p}\cdot\hat{r})\frac{\partial E(r)}{\partial r} \\
&\approx q\,d\,E_o\frac{\partial}{\partial r}\left(e^{-\frac{r^2}{2\sigma^2}}\right) \\
&= -q\,d\,E_o\frac{r}{\sigma^2}e^{-\frac{r^2}{2\sigma^2}}.
\end{aligned}
\tag{5.4}
$$

The electric field E_o where the tweezing force is at its peak can be estimated from

the energy density (peak beam power, P) using:

$$E_o = \sqrt{\frac{2P}{n}\sqrt{\frac{\mu_o}{\epsilon_o}}} \times \frac{\text{measured } E(t \sim 1ps)}{\text{measured } E(t \sim 0ps)}, \tag{5.5}$$

where $\sqrt{\frac{\mu_o}{\epsilon_o}} = 377$ ohms, $n \approx 3.5$, $E(t)$ is the summed fields of the probe pulse and the radiated polarization (without pump excitation) measured by cross correlating the transmitted probe and LO, and the peak in E corresponds to the maximum field strength of the probe pulse. Though we use the nonlinear polarization measurement to determine the form of the tweezing field, we have to use the linear interference term to quantify the peak field strength. Since the macroscopic polarization is much weaker than the pulse (shown in Fig. 5.2) we must consider this ratio when relating the field strength and total excitation power.

The calculated force at the peak of the coherence is then divided by the total mass of the exciton, which is approximately $0.42m_o$, to find the peak in the acceleration.

The model we present considers drag and diffusion. Diffusivity constant, D, for a exciton in a GaAs quantum well should be between 40 and 100 cm^2/s depending on the lattice temperature [95]. We, however, take enough data points in each scan to approximate this.

5.8 Outlook

By introducing the large polarizability of quasiparticle excitons at the resonance frequency of the dielectric medium in which they are created, this work opens the door for many further experiments. Though we measure spot-size dynamics that seem correlated with the radiating polarization, in principle it is possible to use an additional spectrally narrowed pulse to actively tweeze the excitonic polarization. Alternatively, the proposal to create a tweezing potential with repulsive gradient beams [88] may still be implemented, though we expect a smaller effect for similar

fields since the real part of the electric susceptibility peaks above the resonance. Beyond the use of an external beam, the magnitude of the effect will be increased by increasing the optical gradient. This may be accomplished by a near-field tweezing source or an optical mask to create a near-field light source at the quantum well.

This new trapping force on quasiparticles is analogous to that used for optical traps in atomic physics, and so the applications of such a trap are well understood. The introduction of electrostatic traps in semiconductors led to a long interest in condensation of indirect excitons in these traps [96]. Optical gradient traps may have the potential to create condensates of direct excitons. The measurements presented here not only have great theoretical consequences for understanding semiconductors in the presence of spatial inhomogeneity, but must necessarily be considered in experiments with strong field localization by a near-field probe.

CHAPTER VI

Inducing Coherent Quantum Dot Interactions

This chapter adapted from: Eric W. Martin and Steven T. Cundiff. Inducing coherent quantum dot interactions. *Phys. Rev. B*, 97:081301, Feb 2018

6.1 Introduction

QDs are often described as being non-interacting artificial atoms. Some optical spectroscopic experiments have been used to claim that there are no measurable many-body effects present for resonant excitation of interfacial QDs, which would support treating these QDs as non-interacting [98]. However, other optical techniques have been used to measure signatures of interactions between these QDs [35, 63, 99]. Outside of the spectroscopic differences, there are discrepancies that exist regarding the presence of many-body effects in QD lasers [10]. The benefits of QD lasers arise from their discrete and narrow energy levels, but they are usually pumped by the excitation of delocalized carriers [100]. Since many-body effects play a tremendous role in the theoretical treatment of semiconductors [9], it is important to understand the relevant interactions for calculating QD laser properties.

Excitons and trions confined to QDs are potential candidates for qubits in a quantum computer [101, 102, 103]. The electronic states of a QD are easily accessible both optically and electronically. Also, the high oscillator strengths of electronic transi-

tions in solid state systems facilitate their measurement and manipulation. Coherent control with ultrafast Rabi rotations has been demonstrated on both single and ensemble QD systems [104, 105]. However, controlled qubit interaction remains one of the most challenging requirements of a functional quantum computer with few implementations for spin states in QDs [106, 107] and none for the electronic states. The localization of excitons in QDs that gives them the benefit of being difficult to decohere also makes them difficult to entangle, or couple [108].

Here we measure that the excitation of delocalized states not only enhances many-body effects, in agreement with theory [109], but can also turn them on. The physical mechanism responsible for enhancing many-body interactions in QDs may explain the discrepancies in the literature. The mechanism may also be applied for turning on electronic coupling between isolated QD states.

6.2 Experimental Methods

We use ultrafast coherent spectroscopy techniques to directly probe coupling and many-body interactions in a sub-micron-sized region containing a small number of distinct epitaxially-grown GaAs interfacial QDs that are cooled to a temperature of 6 K. These interfacial QDs are exciton states bound by monolayer fluctuations in a narrow 4.2 nm GaAs quantum well with $Al_{0.3}Ga_{0.7}As$ barriers [60]. The decreased transverse confinement binds the localized excitons by 10 meV, which energetically separates them from the delocalized quantum well resonances. Because of the large spatial separation (averaging 300-400 nm) between QDs, the natural coupling between them is minimal. By resonantly exciting higher energy delocalized exciton states in the quantum well we open coherent coupling channels between localized excitons. After pulsed excitation of the delocalized states, we use double-quantum spectroscopy to directly measure exciton-exciton interactions between isolated single quantum systems.

To probe the localized QD response to resonant excitation of delocalized states, we use multidimensional coherent spectroscopy (MDCS). MDCS is a transient four-wave mixing spectroscopy that has evolved from four-wave mixing techniques responsible for realizing the importance of considering Coulomb interaction effects in semiconductor quantum wells [44, 45]. In MDCS, the phase-resolved evolution of an induced nonlinear response is measured as a function of the evolution of a phase-resolved linear response. These measurements result in spectra with two or more dimensions that correlate absorption, emission, and evolution energies of sample coherences [110]. There are various pulse sequences we can use to measure coherent processes. Single-quantum pulse sequences developed for MDCS techniques are used here to directly measure coupling between QD states and the intrinsic linewidth of the QD resonances. A double-quantum pulse sequence is used to directly measure signals resulting from many-body interactions.

Most MDCS techniques rely on k-vector selection, which requires a finite spot typically between 15 and 50 microns. With few exceptions [13, 26, 35], these techniques are thus limited to the study of spatially extended states or dense ensembles. We have developed a variant of collinear techniques [13, 26] that instead uses heterodyne detection to measure radiated MDCS signals. To distinguish the optical signal from the co-propagating excitation beams, each beam is tagged with a different radio frequency using acousto-optic modulators. The radiated third-order nonlinear signals are shifted by radio frequencies that depend on the excitation beams used to generate them. The interference between the radiated nonlinear signal and a separately tagged local oscillator (LO) thus has a beat note at the difference between their frequency tags. A lock-in amplifier is used to measure the signal at the phase-matched modulation frequency. We acquire an accurate measurement of the signal phase by co-propagating all beams with a continuous-wave (CW) laser that samples all of the mechanical fluctuations that contribute to phase noise. We interfere these CW beams

with each other on two detectors and use those measurements to calculate a phase-corrected reference at the signal frequency. Since the signal reference is affected by the same path-length fluctuations as the signal, the measured signal has a meaningful phase with respect to the excitation pulses. See the Chapter II for details about the MDCS.

The MDCS pulses are spectrally tuned and filtered to resonantly excite only the localized exciton states. In Figure 6.1 we compare a single-quantum nonlinear MDCS measurement to the photoluminescence spectrum from the same sample region. These spectra do not need to match because nonlinear emission and radiative recombination for a resonance have different dependences on each QD's dipole moment. In this case, however, the strong localized state resonances, labelled 1-4, emit at the same center frequencies with mostly comparable strengths. Exceptions exist at higher energies, and here we show that resonance 4 has an exceptionally high nonlinear response relative to the lower energy resonances. Since these higher energy states are generally more delocalized, we attribute this enhanced nonlinearity to many-body effects, which are known to be the dominant source of nonlinear optical responses in semiconductor quantum wells [7, 44, 45].

Using a rephasing pulse sequence, which is typically used in ensemble MDCS measurements to separate inhomogeneous and homogeneous broadening such that homogeneous linewidths can be measured [41], we measure an average low temperature QD linewidth between 27 and 28 μeV. This measurement is in agreement with previous low excitation density experiments [60, 99, 111]. At high excitation densities there has been some disagreement in linewidth measurements found in the literature. Four-wave mixing measurements of interfacial QD ensembles have observed large dephasing rates at high densities, a feature that would make these QDs resemble higher dimensional systems [63, 99]. However, linewidth measurements of interfacial QDs with high enough spatial resolution to distinguish the QDs do not depend on ex-

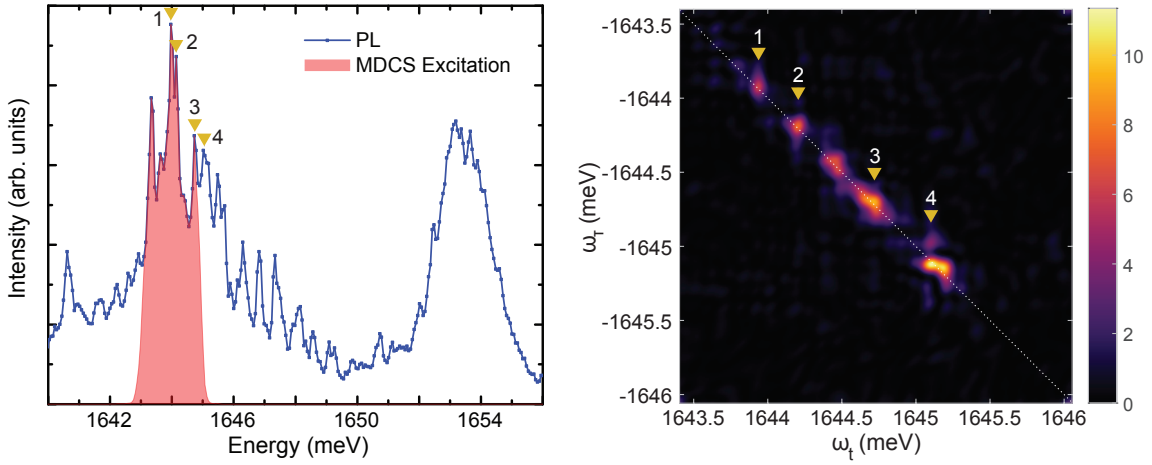


Figure 6.1: Top: Photoluminescence (PL) excited by a 633 nm laser is measured on a spectrometer with 100 μeV resolution. Features below 1650 meV are attributed to localized quantum dot states that we spatially isolated with a diffraction limited 700 nm spot. The wide feature above 1650 meV are the residual two-dimensional (quantum well) states. The region measured by multidimensional coherent spectroscopy (MDCS) in this paper, shaded in red, is determined by the shaped laser spectrum we use. Bottom: Single-quantum MDCS spectrum of the same region allows for comparison of the oscillator strengths of resonances and reveals that some of the weakly excited higher energy states have very high oscillator strengths.

citation density [98]. Taking aspects from all these experiments to understand the source of the disagreement, we use spectrally narrowed pulsed light that only excites localized states and a small excitation spot. We measure that the low temperature linewidth is independent of excitation density and conclude that a likely source of dephasing in experiments with large spot sizes is sample heating.

In order to observe a double-quantum MDCS signal, it is necessary that two interacting excited states coherently evolve simultaneously. This means that incoherent states created by the prepulse (both localized and delocalized) cannot directly induce a double quantum signal. The signal can only be produced by their shifting or dephasing of the doubly excited coherence created by the MDCS beams. With resonant excitation of only the QD states, these signals can result from three interactions. A double-quantum signal resulting from 1) biexcitons in non-interacting self-assembled QDs has been measured [112], but excitation of these signals requires

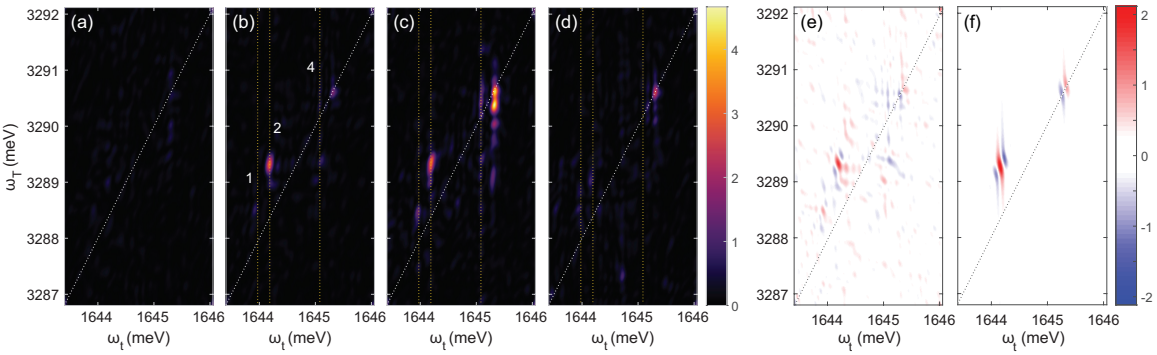


Figure 6.2: Double-quantum MDCS spectra as a function of prepulse power for (a) 0, (b) 500, (c) 1500, and (d) 4000 photons per pulse. A coupling feature between QDs 2 and 4 appears in (b), which corresponds to a many-body interaction that has been turned on between those resonances. A new higher energy feature grows to a maximum on the diagonal at 1645.3 meV in (c). In (d) the prepulse has saturated the system so the coherent signal is significantly degraded. The real part of (b) is plotted in (e), and a simulation of these features is plotted in (f). The simulation is used to determine the dominant many-body terms that give rise to each of the double-quantum signals.

enough bandwidth to excite both the exciton and biexciton. Our sample has been well characterized using MDCS, and it is known that the biexciton binding energy of an ensemble of these dots increases with emission energy from 3.3 to 3.8 meV, and it has a distribution about that center binding energy of less than 270 μeV [113]. The distribution of biexciton binding energies for a set of individually measured quantum dots, which has the advantage over ensemble measurements of being able to exactly correlate biexciton and exciton emissions, is just 200 μeV [114]. The MDCS beams have a narrow bandwidth of 2 meV with sharp spectral edges (0.2 meV) such that we cannot doubly excite a single QD. Thus, the only source of a signal from a QD resonance can be 2) interaction between two different QDs. The interactions in both measurements, however, are very weak and require that the QDs have a very close proximity. 3) If a weakly localized state is large enough for it to be doubly excited without forming a bound state, the resulting double-quantum signal would more closely resemble those measured in quantum wells [115]. Using the phase of the double-quantum signal, we can distinguish binding and scattering many-body

interactions [116], so we can identify the above sources of double-quantum signals. By spatially isolating just a few quantum dots within a 700 nm focus, we can thus directly measure interactions at the single excitation level.

We use double-quantum MDCS to determine if QD states produce interaction induced signals. On-diagonal signals in a double-quantum MDCS spectrum correspond to self interaction, which we attribute to either spatially large localization sites that confine multiple degenerate excitations or spatially adjacent nearly degenerate quantum dot sites. Off-diagonal signals are due to many-body interaction between two excitations at different energies. For example interactions between two frequencies ω_1 and ω_2, these signals can emit at either of those frequencies and will evolve at their sum: $\omega_T = \omega_1 + \omega_2$. These weak off-diagonal signals most likely result from radiative interaction between adjacent quantum dot states, which has been shown to have a long range exceeding 400 nm [35]. Though weak, we measure interactions between few resonantly excited QDs over the sample, and weak interactions between resolved QD states have very recently been measured in self-assembled QDs [112].

6.3 Results

In order to measure the effect of delocalized quantum well excitations on QD interactions we excite the delocalized quantum well states with a pre-pulse that is spectrally filtered to excite only the quantum well states. The pre-pulse has a power between 10 and 80 nW (500-4000 photons per pulse), and it arrives 20 ps before the first MDCS pulse so that only the incoherent population it creates is present when the MDCS spectrum is measured. As shown by comparison of Figs. 6.2(a) and (b) a small excitation of the delocalized states greatly enhances interaction among localized QD excitons, which results in a strongly enhanced off-diagonal peak in the double-quantum MDCS spectrum. From the evolution and emission energies it is evident that this feature corresponds to coupling between resonances 2 and 4 labeled

in Fig. 6.1. As the prepulse power is increased in Fig. 6.2(c), lower energy QD states are filled due to dynamic localization of the extended states created by the prepulse, and higher energy double-quantum features are enhanced. The strong on-diagonal feature at $\omega_t = 1645.3$ meV does not strongly show up in single-quantum MDCS without some prepulse excitation either, shown in section 6.5. Along with the enhanced oscillator strength of the high energy features measured with single-quantum MDCS, the presence of this state on the diagonal is evidence that it is a higher dimensional state than a QD since it can be doubly excited. With a much higher prepulse excitation in Fig. 6.2(d), all double-quantum coherences are blocked by filling of the QD states.

In Fig. 6.2(e) we plot the real part of (b). To interpret the phase of the double-quantum MDCS signal requires a simple simulation in which we consider the phase of the linear responses to each pulse. We simulate the nonlinear response by analytically solving a perturbative expansion of the density matrix for two coupled two-level systems [94]. The energy level scheme consists of a ground state, a singly excited state for each QD, and doubly excited state representing simultaneous excitation of both QDs. Many-body interactions break the symmetry between the transition into the singly excited and doubly excited states, which is represented by a shift or broadening of the doubly excited level. By accurately measuring the phase of the double-quantum signal, we can identify the many-body terms that give rise to those signals. In order to produce an accurate simulation of the data in Fig. 6.2(f), we find that the coupling feature corresponding to the interaction of QDs 2 and 4 is an excitation induced red shift of the doubly excited state. A red shift of the doubly excited state is indicative of a weak binding between the two states [116]. The on-diagonal feature, on the other hand, results from a combination of excitation induced dephasing and blue shift. These exciton scattering effects are typically measured in quantum wells, further supporting that this higher energy state is higher dimensional

than a QD. We find similar results for QD states in other regions of the sample.

6.4 Conclusions

The prepulse enhancement of many-body interactions between QD states is illustrated in Figure 6.3. The delocalized carriers in the quasi-continuum states serve to mediate interactions between spatially separated QD states. The QD excitations are localized to roughly 50 nm islands, but delocalized excitons are much more extended. So while there is no wave function overlap of individual QDs, the wave function of the quantum well excitation introduces coupling of localized states. The enhanced range of interaction between QDs is still limited by the finite mobility of the delocalized excitons, roughly 15 cm^2/s in a thin quantum well [54]. Therefore only a few of the localized excitations within a given spot will be within the range of each other to interact via the delocalized excitons.

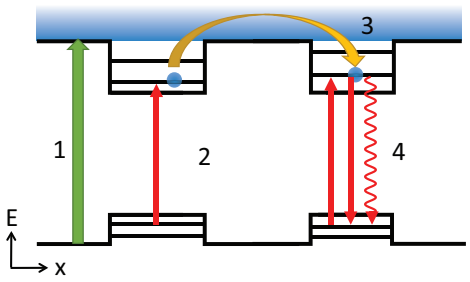


Figure 6.3: Schematic of pulse sequence applied to spatially isolated interfacial quantum dots. Interactions between QDs are very weak, but we can turn on coupling by creating delocalized quantum well carriers with (1) a higher energy prepulse. We probe the induced interactions with (2) two coherent pulses that create a double coherence of different excitonic transitions. (3) The interaction between the coherences is mediated by the quantum well carriers, and (4) we read out the interaction with a coherent third pulse that begins emission of a coherent four-wave-mixing signal. Though the prepulse also creates incoherent excitations of the QD states, this is negligible for low prepulse powers and only serves to degrade the overall signal.

Existing microscopic theory supports the concept that excitation induced dephasing and shift in interfacial QDs arises from interactions with quasi-continuum quantum-well states [109]. Schneider *et al.* discuss broadening and redshift that is

dependent on density, effects they determine by calculating the renormalized electronic states. They also discuss that their calculation of density dependent dephasing in interfacial QDs is equivalently relevant to self-assembled quantum dots electronically coupled to a wetting layer.

Though we have presented a method for turning on coupling between isolated interfacial QDs, this method may be generalized to coupling any localized quantum states in physical contact with higher energy delocalized states; at least states that may be excited in a controlled way. We see immediate benefit in the ability to control coupling self-assembled QDs in contact with the higher energy wetting layer. Also, in light of recent findings of long-lived localized states in transition metal dichalcogenides (TMDCs) [117, 118, 119, 120, 121], this work could be applied to these states which could be coupled through the highly delocalized TMDC exciton states.

Another major outlook for measuring physical systems at the nanoscale is that the coupling of individual QDs to delocalized excitons introduces a new method for studying the delocalized states. The locations of QDs can be determined with much higher accuracy than the optical resolution. Since QDs are spectrally distinct, one could thus consider using measurements of QDs separated by known distances to probe length scales and transport in the continuum states with the resolution of a QD.

In summary, we have developed a collinear MDCS technique that utilizes dynamic phase cycling to probe nonlinear responses with high sensitivity and phase resolution. We have used this technique at the diffraction limit to resolve individual QD oscillators. We demonstrated both double-quantum and single-quantum measurements, and with this technique it is actually simple to selectively measure even higher order nonlinear expansion terms. Using double-quantum MDCS, which is sensitive only to many-body effects, we measure an absence of many-body effects in interfacial quantum dots with resonant excitation. However, we find that these effects can be

enhanced by excitation of the delocalized quantum-well states using a prepulse. This work helps to explain some discrepancies in the literature in which weak excitation of continuum states with broadband pulses has not been explicitly considered. From an applications standpoint, we present this prepulse technique as a way of turning on coupling between quantum states.

6.5 Single-Quantum MDCS Results

We simultaneously measure rephasing and non-rephasing single-quantum spectra using separate lock-in channels. We find that the single-quantum signals, shown in Figure 6.4, are also somewhat enhanced by prepulse excitation of the quantum well states. We also measure filling of the lower energy states as the peak of the nonlinear signal shifts towards higher energy quantum dots (QDs).

An interesting difference between the rephasing and non-rephasing sequences is also realized in this case of isolating individual QDs with a Gaussian beam. The spatial isolation of QDs by a tight focus, as opposed to a mask over the sample, offers the freedom to spatially scan over the sample. Unfortunately it leads to an inhomogeneous excitation of the QDs. Therefore, the strongest contributors to the signal are those QDs that are located near the spatial center of the beam. There are many more QDs near the edge of the beam that contribute to a background signal in our measurements. This is particularly noticeable for signals created by the rephasing pulse sequence in which all the weakly excited QDs constructively interfere to create a photon echo. Though the strongly excited QDs still dominate the signal, in a rephasing spectrum they sit on top of a background along the diagonal of weakly excited QDs, shown in Fig. 6.4(a)-(d). In the non-rephasing spectra, shown in Fig. 6.4(e)-(h) the strongly excited dots are more emphasized as many of the background excitations dephase each other.

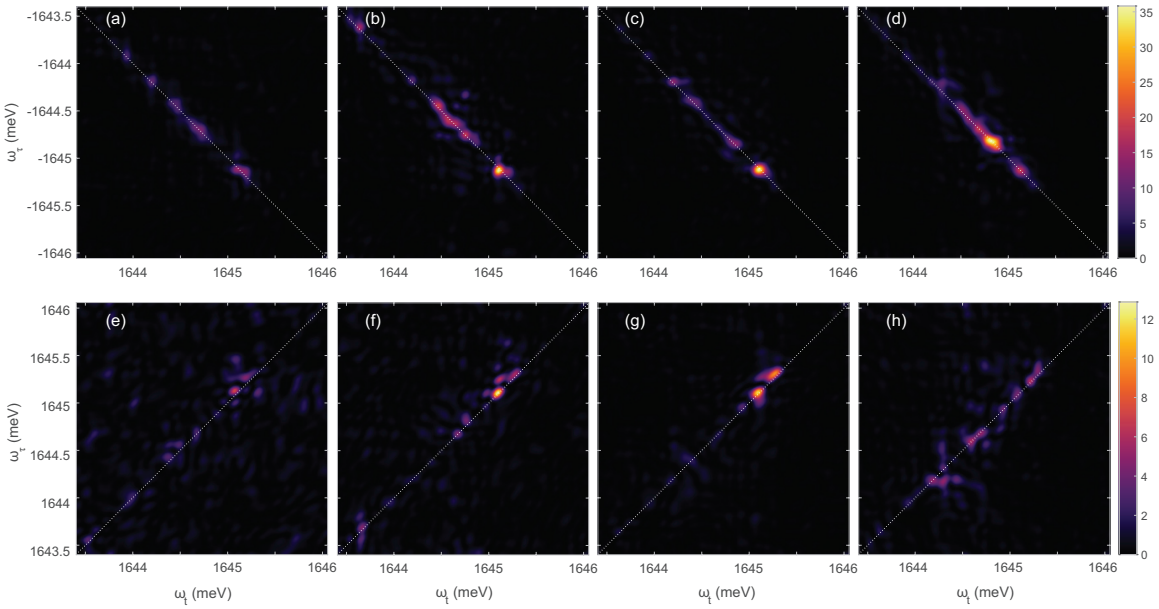


Figure 6.4: Single-quantum spectra are also enhanced by prepulse excitatation. (a) is the rephasing spectrum with no prepulse excitation that is plotted in the paper. The other figures on the top row are rephasing spectra measured with prepulse powers of (b) 500, (c) 1500, and (d) 4000 photons per pulse. (e)-(h) are the corresponding non-rephasing spectra. The complete dephasing of the non-rephasing signal with high prepulse excitation is expected, while the increased rephasing signal strength can be attributed to the enhancement of the weakly excited background excitations.

6.6 Evidence that Continuum Coupled QDs are Spatially Distinct

In a single-quantum measurement, off-diagonal features result from biexciton excitation and interexciton coupling. The biexciton states can be clearly identified by their binding energy shift from the absorption resonance, and all other features result from coupling of spatially distinct exciton resonances [114]. Single-quantum measurements are sensitive to all the interactions that a double-quantum measurement can measure (and more), and therefore the coupling features we measure with a double-quantum measurement will also only result from spatially distinct resonances.

To further support the statement that the prepulse is enhancing coupling between spatially distinct QD resonances, we measure a coupling feature at a different spot on the sample. These data are plotted in Fig. 6.5. In Fig. 6.5(a) and (b) we plot double-quantum spectra for a 0 and 500 photon prepulse, respectively. This demonstrates

that these resonances respond to the continuum prepulse as discussed in the text. In Fig. 6.5(c) we plot a single-quantum rephasing measurement of the same position on the sample with dotted lines indicating the locations of the coupled resonances. We move the sample less than a spot diameter and measure the single-quantum spectrum of the new spot, plotted in Fig. 6.5(d). In the new spot one feature is enhanced while the other is suppressed, indicating that the two signals are emitted from different locations.

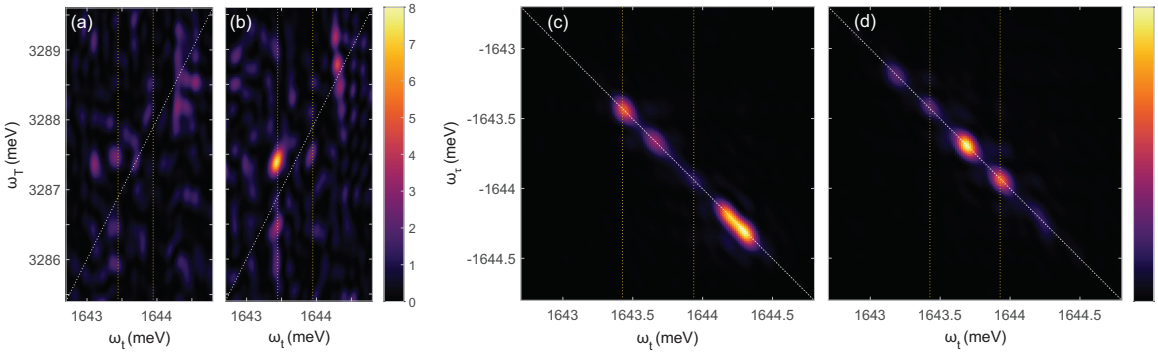


Figure 6.5: We measure an enhanced coupling of two QDs using a prepulse and verify that the two QDs are spatially distinct. We plot double-quantum spectra of a spot on the sample in which we measure a coupling of QDs located at energies indicated by the vertical dotted lines using (a) no prepulse and (b) a 500 photon/pulse prepulse. In (c) we plot a measurement of the single-quantum spectrum at this spot on the sample. In (d) we measure the single-quantum spectrum of a spot spatially shifted less than a spot size away. We see the same resonances in (c) and (d), but the relative intensity of the resonances has changed as the lower energy QD is less efficiently measured by the new spot and the higher energy QD is more efficiently measured by the new spot.

6.7 Sample description

Interfacial QDs are formed by width fluctuations of a GaAs quantum well [60, 61, 62]. The well is epitaxially grown with a nominal width of 15 monolayers, or 4.2 nm. The 35 nm barriers are made of $Al_{0.3}Ga_{0.7}As$. The growth of the top barrier is delayed by tens of seconds to give monolayer width fluctuations of the quantum well time to coalesce into islands. These islands, with a width of 16 monolayers, form the interfacial quantum dots, which are bound by 10 meV. The QDs cover approximately

2% of the sample area and each have a size of approximately 36 nm [63].

6.8 Model of Interactions

QDs can generally be simulated as a simple two-level system. We add an additional weak coupling between two QDs in order to simulate the double-quantum spectra, which results in the energy-level scheme of a diamond configuration plotted in Fig. 6.6. The single excited states, $|e_1\rangle$ and $|e_2\rangle$, are located at 1644.19 meV and 1645.09 meV above the ground state. The doubly excited state is located at 3289.28 meV above the ground state with additional shifting and/or dephasing added as the source of the interaction signal. Since experimentally all of these energies are measured with respect to the evolution of a continuous-wave laser at 1627.5 meV, we also perform our simulations in this rotating frame. In Fig. 6.7 we plot the response of each many-body effect on the real part of the simulated system, and we demonstrate how we can uniquely identify that the doubly excited state is shifted or dephased with respect to the sum of single excited states using double-quantum multidimensional spectroscopy.

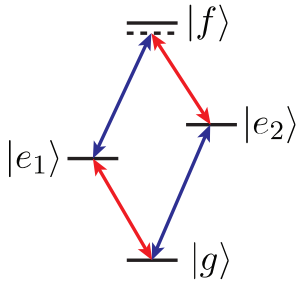


Figure 6.6: The diamond configuration energy level diagram is used to model two coupled QDs with singly excited energy states $|e_1\rangle$ and $|e_2\rangle$. The doubly excited level is labeled $|f\rangle$ with the dotted line indicating that this level is either shifted or dephased from the non-interacting doubly excited level.

In Fig. 6.7, the phase of the signal resulting from a small (with respect to the linewidth of the resonance) excitation induced shift (EIS) of the double excited state is $\pm\pi/2$ from a signal resulting from purely excitation induced dephasing (EID) depending on whether the shift is blue or red. Any possible phase of the double quantum

signal can therefore be determined by a combination of these effects. There is also an overall phase of the signal at the resonance which is uniquely determined by its double-quantum energy and the fixed τ delay at which the measurement is performed. This phase results from the evolution of the resonances during τ, which is the lower energy state in half of the terms and the higher energy state in the other half with an equal oscillator strength for both. The overall phase, ϕ, for any interaction resonance in a double quantum spectrum that can be modeled using coupled two-level systems is $-(\omega_1 + \omega_2)/\hbar \times \tau = -\omega_T/\hbar \times \tau$. If τ is set to 0 in the experiment, one can read off the phase of the interaction signal from the plot to determine EID (0), red EIS ($+\pi/2$), or blue EIS ($-\pi/2$). We typically set τ to a small finite value, however, to guarantee a correct pulse ordering in time. For the coupling resonance shown in the main text we calculate $\phi = -1.82$ rad. The phase of the signal in the experimental data is -0.18 ± 0.19 rad, corresponding to an almost exactly $+\pi/2$ phase shift.

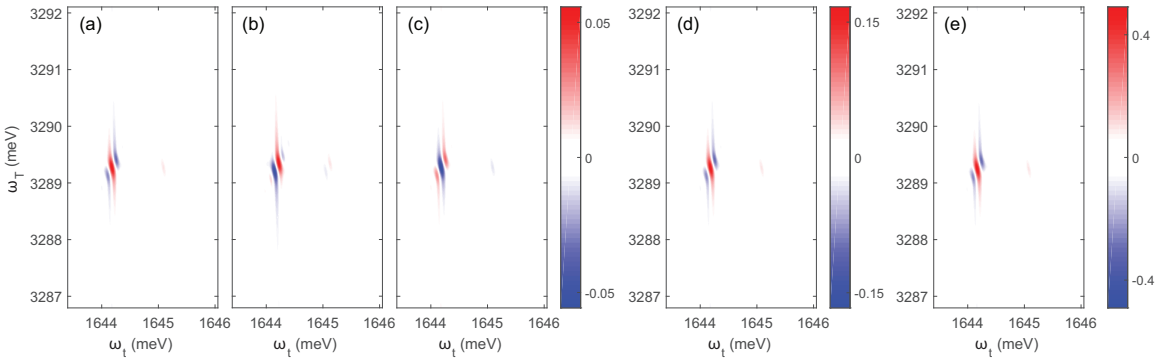


Figure 6.7: We plot simulations of a double quantum signal resulting from a small (a) excitation induced shift of -0.003 meV, (b) excitation induced dephasing of 0.003 meV, and (c) excitation induced shift of +0.003 meV. We also show that as the magnitude of the excitation induced shift is increased, but still below the 0.05 meV linewidth, the general shape of the feature is unaffected. For (d) a shift of -0.009 meV and (e) a shift of -0.027 meV, the amplitude of the signal increases, but the phase does not.

Simulation of the diamond level structure requires the response function of all possible third-order nonlinear terms be considered. These response functions can be diagramatically represented with double-sided Feynman diagrams [94, 116]. In

Fig. 6.8 we show the diagram corresponding to terms for which the sample is initialized in the ground state. There are four equivalent diagrams in which the first absorption is from $|g\rangle$ into the $|e_2\rangle$ state. Since we fix the τ delay, we do not resolve this. One can see by looking at these double-sided Feynman diagrams that excited state emission diagrams 1a and 2a have an opposite sign of the excited state absorption diagrams 1b and 2b. Also, without any interaction induced effects between $|e_1\rangle$ and $|e_2\rangle$, the emission energy of diagrams 1a and 1b (2a and 2b) is the same. Therefore, one can only measure a double-quantum signal if the emission energy from the doubly excited $|f\rangle$ to $|e_2\rangle$ differs from $|e_1\rangle$ to $|g\rangle$. This can only result from an interaction induced shift or dephasing of the doubly excited state [116].

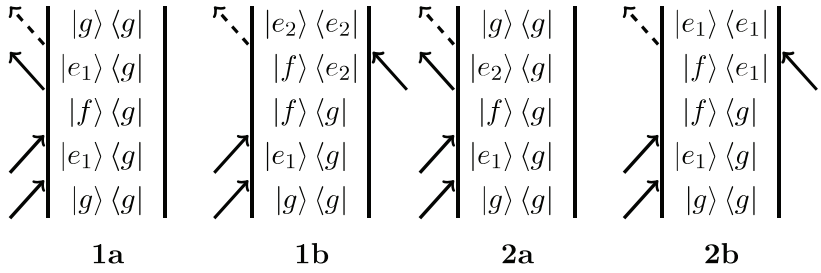


Figure 6.8: These are double-sided Feynman diagrams corresponding to excitation of a double-quantum signal in the diamond configuration energy-level structure. These are all initialized in the ground state with the first excitation into the first excited state $|e_1\rangle$(there are four more equivalent diagrams first excited into $|e_2\rangle$). 1a and 1b both emit at the frequency corresponding to $|e_1\rangle$ and have opposite signs. A signal is emitted because there is some energy shift or dephasing of the double excited state $|f\rangle$ that is due to interaction between the two coherently excited states. 2a and 2b both emit at the $|e_2\rangle$ frequency and also have opposite signs.

In Fig. 6.9 we plot diagrams in which the prepulse has unintentionally excited an incoherent population in the localized states. We show diagrams for a singly excited population state and doubly excited population. Diagrams 1c-f give an identical lineshape to diagrams 1a and 1b, except that for the same emission energies they have an opposite sign. This means that no new signals arise from the introduction of incoherent localized states, and as with Fig. 6.8 the signal results from a shifting or dephasing of the $|f\rangle$ state in Fig. 6.6 by the prepulse excitation. There are also twice

as many diagrams for the prepulse excitation into the single excited state. For initial excitation of both involved quauntum dots, the sign is flipped again for constructive interference with the ground-state-initialized diagrams. Since $|\langle e_1 \,|\, e_1\rangle|^2 < 0.5$, these diagrams will not completely cancel. Thus measurements of the signal phase can be used to uniquely determine the relative energy shift and dephasing of the excited state.

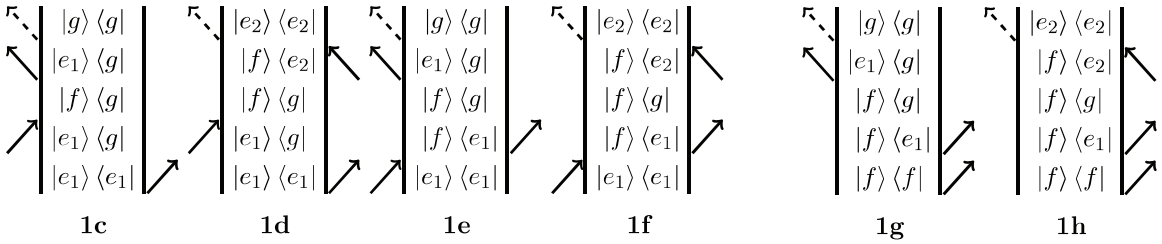


Figure 6.9: Though the prepulse is resonant with the quasi-continuum states, some of those will relax into localized states prior to measurement by coherent spectroscopy. We must therefore consider initialization into the singly (1c-f) and doubly (1g-h) excited population states. The shapes of these response functions actually all look the same as the ground state diagrams. The signs of the singly excited states are all opposite, however, and therefore degrade the magnitude of the double quantum signal without distorting the phase. Here we only plot a fourth of the total number of diagrams, leaving out nearly equivalent diagrams and those with emission at the $|e_2\rangle$ frequency.

CHAPTER VII

Encapsulation Narrows Excitonic Homogeneous Linewidth of Exfoliated $MoSe_2$ Monolayer

7.1 Introduction

Monolayer van der Waals crystals are a class of materials with widely varying properties and the potential to transform future electronics [122, 123, 124]. These atomically thin layered materials can be stacked into heterostructures with synergistic benefits [125]. A subset of these materials are the semiconducting monolayer transition metal dichalcogenides (TMDCs), which have direct band gaps that make their electronic transitions optically accessible and thus useful for optoelectronic applications [126, 127, 128, 129]. The low dimensionality resulting from confinement to a monolayer also means monolayer TMDCs have very strong many-body interactions that result in $\sim$100 times larger binding energy of excitons than more conventional III-V semiconductors, like gallium arsenide. Excitons thus dominate the optical response of semiconductor TMDCs and remain strongly bound at room temperature. The low dimensionality also means excitations of these materials are very sensitive to the external environment through both screening and introduction of defects.

Encapsulation of monolayer van der Waals crystals by hBN has proved to enhance carrier mobility [130, 131] and significantly improve the monolayer resistance to photodegradation [132]. Most notably hBN encapsulation has been shown to

greatly reduce the photoluminescence linewidth of $MoSe_2$, MoS_2, WSe_2, and WS_2 [133, 134, 135].

Narrowing of the photoluminescence linewidth of a TMDC monolayer is an indicator that encapsulation passivates the monolayer surface and minimizes inhomogeneity resulting from trapped states and defects. However, narrowing of the photoluminescence linewidth can also result from a change in the radiative linewidth of the exciton, which scales with the substrate index. The substrate can further affect the linewidth through its effect on pure dephasing resulting from interactions of the exciton with photons, phonons, and other collective modes. So while it is impressive that encapsulated TMDC photoluminescence linewidths approach the homogeneous limits measured in similar monolayers on different substrates, more direct measurements of the various excitonic linewidths are necessary to fully determine the effect of encapsulation. Specifically we must disentangle the linewidth contributions from inhomogeneous broadening, spectral diffusion, and radiative decay.

Homogeneous linewidths of bare samples have been measured using four-wave-mixing (FWM) based techniques [136, 137, 138]. Multidimensional coherent spectroscopy (MDCS) was employed by Moody *et al.* to measure the homogeneous linewidth of WSe_2 grown by chemical vapor deposition (CVD) [136] and to identify higher order correlated states in a large exfoliated $MoSe_2$ flake [139]. MDCS is useful for its ability to unambiguously separate homogeneous and inhomogeneous broadening of exciton linewidths. However, their implementation of MDCS was limited by having a relatively large spot size ($\sim$30 μm) requiring large samples, and CVD grown samples are known to show poor quality and mid-gap defects as compared to exfoliated samples [140, 141]. Jakubczyk *et al.* used three-pulse FWM microspectroscopy to measure exfoliated $MoSe_2$. This FWM technique provides similar information to MDCS for lineshapes without inhomogeneous dephasing rates [43]. The authors demonstrate tremendous variability of the exciton transition energy (greater than 10

meV) and dephasing time (between 0.5 and 1.5 ps) over the single large exfoliated flake [137].

7.2 Determining Linewidth Contributions

Here we use MDCS in conjunction with linear reflectance spectroscopy to compare the neutral exciton linewidths of fully hBN-encapsulated and non-encapsulated exfoliated $MoSe_2$ monolayer samples. Prototypical samples, including heterostructures consisting of exfoliated monolayer materials, are often small. The samples measured here are all between 5 and 8 μm wide. Example images of two of these samples are presented in Fig. 7.1(a). The small sample size demands the use of collinear techniques for optical measurements. For linear spectroscopies including linear absorption, plotted in Fig. 7.1(b), the excitation source is easily distinguished from the sample response. For MDCS it is necessary to distinguish a specific third order response from the excitation sources, linear response, and all other nonlinear responses. The conventional method for isolating the MDCS signal is with wave vector selection [110], but this is not congruent with having a collinear technique. We have developed a collinear MDCS that has enabled measurements of samples within a diffraction limited spot [97] based on frequency modulation [13].

Using MDCS we are able to extract the energy dependent homogeneous linewidth in the presence of inhomogeneity. The homogeneous linewidth $\gamma = (\Gamma_{rad}+\Gamma_{SD})/2+\gamma^*$, where Γ_{rad} is the radiative decay rate, Γ_{SD} is the spectral diffusion rate, and γ^* is the broadening due to pure dephasing. The spectral diffusion rate accounts for inelastic scattering into other excited states and localized states [54]. Typical pure dephasing processes are elastic exciton-exciton and exciton-phonon scattering. A nearly exhaustive list of decay processes in semiconducting TMDCs are detailed by Moody *et al.* [142], but at very low temperature these are summarized by relaxation to lower energy states. By careful measurement and sample selection we can determine and suppress

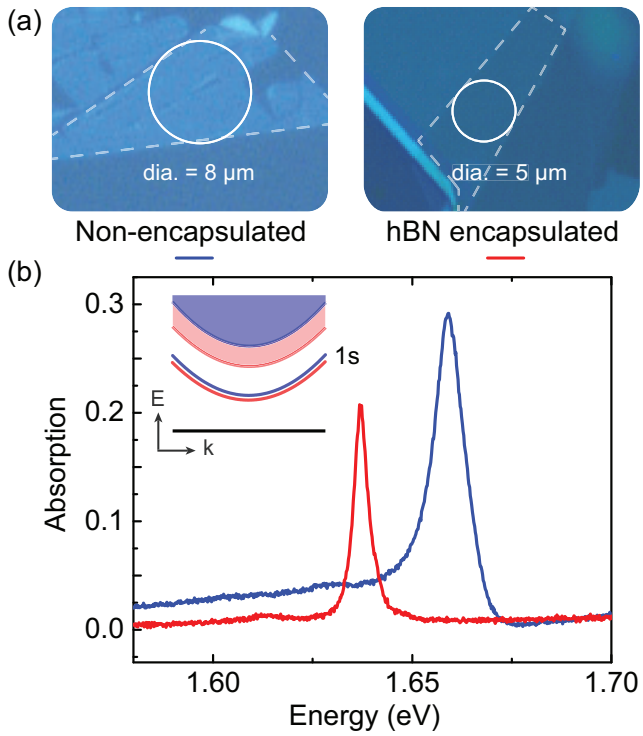


Figure 7.1: (a) Microscope images of non-encapsulated (left) and hBN encapsulated (right) exfoliated $MoSe_2$ monolayer samples. These images illustrate the fairly small size of the measured samples. (b) Linear absorption spectra of these two samples, calculated using linear reflectance spectra. The encapsulated sample has a decreased linewidth and decreased intergrated absorption, which are indicators that the sample inhomogeneity has been reduced and spectral diffusion processes have been minimized. (inset) Encapsulation decreases the band gap and exciton binding energy so that the transition energy to the exciton of the encapsulated samples (red line with continuum shaded red) is about 20 meV lower than in the non-encapsulated samples (blue line with continuum shaded blue).

the majority of these decay and dephasing terms. Exciton-exciton interactions are excitation density dependent, and we measure the excitation density dependence of the linewidth to determine their contribution to γ. Exciton-phonon scattering to higher energy states can be suppressed by lowering the sample temperature to nearly 0 K. At low temperatures the phonon broadening is due to acoustic phonons, and we measure the temperature dependence of the linewidth between 5 and 80 K to determine their contribution to γ.

The low temperature contributions to Γ_{SD} by decay to lower energy states are

more difficult to determine by measurement. The presence of indirect exciton states, charge defects, and localized states all dephase the direct neutral exciton even at low temperature. These three states all have sample dependences that we can consider, however. 1) Recent experimental and theoretical studies have demonstrated the existence of dark indirect excitons at a lower energy than the bright exciton in tungsten-based materials [143, 144, 145]. Measuring molybdenum-based $MoSe_2$ and MoS_2 avoids this complication. 2) hBN encapsulation of monolayer TMDCs has proved a method for minimizing charge defects affecting the monolayer, evidenced by the suppression of the trion resonance in encapsulated $MoSe_2$ [134]. 3) Strongly localized defect states are absent in most samples on sapphire, and their absence can be easily confirmed by photoluminescence.

7.3 Linewidth Comparison of Encapsulated and Non-encapsulated $MoSe_2$

Here we compare four high quality samples: two monolayer $MoSe_2$ mounted on sapphire substrates and two hBN encapuslated monolayer $MoSe_2$ also mounted on sapphire. The encapsulated monolayers are on a bottom layer of hBN that is approximately 120 nm and a top layer that is approximately 20 nm, measured with atomic force microscopy. In Fig. 7.1(b) we plot linear absorption spectra of one of the non-encapsulated and one of the encapsulated samples to demonstrate that their linewidths are comparable to the best observed in the literature [146, 147]. The significant decrease of the total absorption of light by the encapsulated sample is an additional indicator that encapsulation decreases the scattering processes that contribute to incoherent absorption. We also see that the transition energies of excitons in the two samples differs by about 20 meV. This is primarily due to the significant decrease of both the band gap and partially compensating exciton binding energy by encapsulating the monolayer in a high index material [148, 149]. We depict these changes in the niset of Fig. 7.1(b). With MDCS we observe that encapsulation signifi-

cantly narrows both the inhomogeneous and homogeneous exciton linewidths, and the homogeneous linewidth variance is much lower for encapsulated samples. We discuss how this implies that hBN-encapsulation of monolayer TMDC samples minimizes defects and static doping that result in both long- and short-range disorder. Along with the static lineshape differences, we measure significant permanent modification of the homogeneous linewidth of non-encapsulated samples resulting from temperature cycling and exposure to weak radiation. In contrast, encapsulated samples are very robust to numerous temperature cycles and high radiation exposure.

In Fig. 7.2(a) we plot characteristic multidimensional coherent spectra of low temperature samples using a rephasing pulse sequence. To generate these plots we measure the phase-resolved evolution of an induced nonlinear response as a function of the evolution of a phase-resolved linear response. We measure these responses using a sequence of four pulses in the time domain having relative delays (t, T, and τ) that are referenced to a co-propagating continuous-wave laser. Fourier transforming the response with respect to the pulse delays yields spectra with multiple dimensions that correlates absorption, emission, and evolution energies of the sample coherences [97, 110]. Specifically we use a pulse sequence for correlating absorption (ω_τ) and emission (ω_t) energies. The evolution of the absorption in a rephasing measurement has the opposite sign to the emission, so these frequencies are negative. Since a single resonance absorbs and emits at the same energy, a distribution of resonances will all fall along the diagonal axis where $-\omega_\tau = \omega_t$. Thus the lineshape of the diagonal slice plotted on the left in Fig. 7.2(b) roughly corresponds to inhomogenous distributions of exciton resonances in the non-encapsulated (blue) and hBN-encapsulated (red) samples. The lineshape of the cross-diagonal slice plotted on the right roughly corresponds to the homogeneous linewidth of those exciton resonances. For more a more detailed description of the MDCS experiment, see Chapter II.

Though the cross-diagonal linewidth roughly corresponds to the homogeneous

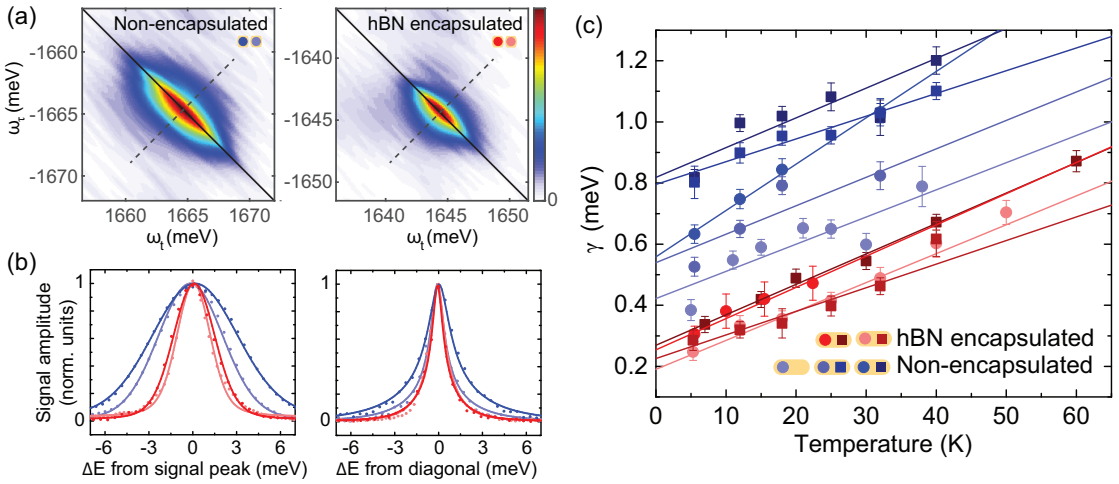


Figure 7.2: (a) Characteristic low-temperature, low-power multidimensional coherent multidimensional spectra of non-encapsulated $MoSe_2$ on sapphire and hBN-encapsulated $MoSe_2$. (b) Slices along the diagonal (left) of a multidimensional spectrum roughly correspond to the inhomogeneous distribution of exciton resonances. Slices along the cross-diagonal (right) roughly correspond to the homogeneous lineshape. We plot these slices for low temperature, low power measurements of four samples: two non-encapsulated samples in blue and two encapsulated samples in red. Since the diagonal and cross-diagonal slices are correlated, it is essential to fit them simultaneously to determine the homogeneous and inhomogeneous linewidths [3]. (c) We plot extrapolated zero power linewidths of each sample as a function of temperature. Grouped in the legend, circle data points correspond to a first measurement set of the sample and square data points correspond to a measurement set made after temperature cycling the same sample.

linewidth, the lineshape is very dependent on the inhomogeneous distribution. For low inhomogeneity the cross-diagonal lineshape is a Lorentzian. However, in the limit of high inhomogeneity the lineshape is the square root of a Lorentzian [41]. Ignoring inhomogeneous broadening and using the wrong fit function when determining a sample's homogeneous linewidth (or dephasing times in photon echo FWM experiments) can thus significantly skew the measurement, up to a factor of $\sqrt{3}$. For inhomogeneous linewidths that are comparable to the homogeneous linewidth, as is the case in these samples, it is essential to simultaneously fit the co-dependent diagonal and cross-diagonal slices. Here we actually fit the entire two-dimensional spectrum using an analytical solution to the optical Bloch equations (OBEs) derived by Bell *et al.*

[3]. Though the lineshapes here are sufficiently narrow relative to the laser bandwidth to not require it, we fit broader linewidths using an analytical solution to the OBEs that includes finite pulse effects [42].

We measure the homogeneous linewidth as a function of beam power and sample temperature. We increase the power of all three excitation beams equally, and determine the linewidth scaling as a function of the excitation density of a single beam. We estimate the linewidth linearly broadens with a slope of 4×10^{-13} meV/photon·cm^{-2}. We measure this linear dependence up to 10^{12} photon·cm^{-2}. For each temperature we extrapolate the power dependence of the linewidth to zero power and plot that as γ. In Fig. 7.2(c) we plot γ as a function of temperature for four different samples. The non-encapsulated samples are indicated in blue. The circle data points correspond to a first set of measurements on a sample, where the linewidths are first measured at 5.3 K and the temperature is increased. The square data points correspond to measurements made after a temperature cycle defined by warming the sample up to room temperature and cooling back down to again start the measurement set at 5.3 K. It is evident from this data that the linewidth of the non-encapsulated monolayer is very sample dependent, which confirms results by Jakubczyk *et al.* [137, 138]. We further find significant broadening of the exciton linewidth of non-encapsulated samples with a single temperature cycle. By measuring many points on the sample, we confirm that the broadening effect is not the result of a positioning error. We rather suggest that the broadening is likely a result of the deposition of molecules such as water on the sample surface. Whether the change results from surface molecules or substrate strain, we demonstrate that the homogeneous linewidth is a sensitive indicator of a change in the sample environment. The hBN-encapsulated monolayer samples are indicated in red. The sample variance is very small, and there is no measurable broadening due to temperature cycling in these samples. This consistency is evidence that defect scattering is minimal in encapsulated samples. The durability of the

monolayer with temperature cycling is an important confirmation that experiments on encapsulated samples will be consistent and reproducible.

From the temperature dependence we measure a linewidth broadening of 0.010 ± 0.001 meV/K by averaging the hBN encapsulated sample linewidths. This exciton-acoustic phonon interaction is similar to those measured for the non-encapsulated samples confirming previous results [137], similar to the interaction in ZnSe (slope = 8 μeV/K) [150], and approximately double the interaction GaAs quantum wells (slope = 5 μeV/K) [151]. We determine an average extrapolated zero temperature and zero power linewidth, $\gamma_0 = 0.26 \pm 0.02$ meV, for the encapsulated sample. This corresponds to a dephasing time $T_2 = \hbar/\gamma \approx 2.5$ ps. Since the sample variance of the non-encapsulated samples likely results from surface molecules and substrate effects, the broadest linewidths are dominated by scattering processes. Though it is important to note that these measurements can only identify an upper bound of the radiative linewidth, we identify the lowest measured linewidth for the non-encapsulated samples as the nearest to the radiatively limited linewidth for these samples. For this non-encapsulated sample $\gamma_0 = 0.42 \pm 0.05$ meV, which corresponds to $T_2 \approx 1.6$ ps. This latter value is in agreement with previous determinations of the longest T_2 times measured in $MoSe_2$ using FWM [137] and time-resolved photoluminescence [152].

We attribute the difference in homogeneous linewidths between the hBN encapsulated monolayer and the monolayer directly on sapphire to the difference in the dielectric environments. The radiative linewidth should scale with the substrate refractive index: $\gamma_{rad} \propto 1/n_{top,\ bottom}$ [145]. For a radiatively-limited homogeneous linewidth, one would thus expect an hBN encapsulated sample to have a $\frac{n_{sapph}+n_{vac}}{2\times n_{hBN}} \sim 0.7$ times narrower linewidth than an equivalent sample on sapphire. The relative scaling between the encapsulated samples and narrowest non-encapsulated sample is 0.62, which is consistent with the expected narrowing for radiatively-limited linewidths. We thus present this as evidence that we are in fact measuring a nearly radiatively-limited

linewidth and consistently measure minimal exciton scattering in hBN encapsulated samples.

7.4 Photodegredation Effects on $MoSe_2$ Samples

Finally we compare photodegradation of low temperature samples resulting from excitation by resonant pulses. We treat samples by irradiating them with laser light for one minute at the given treatment beam power. The pulsed light is focused to a 2 μm spot and has a repetition frequency of 76 MHz. After each treatment we turn off the treatment beam and measure a multidimensional spectrum with low power, 1 μW/beam, pulses. In Fig. 7.3 we plot the MDCS signal strength as a function of treatment pulse power. We find that the non-encapsulated samples exhibit lasting damage by beams having powers greater than 45 μW. Measured homogeneous linewidths varied significantly between treatments and scans. The encapsulated samples, however, are resilient up to powers that fully saturate the exciton and have a consistent homogeneous linewidth. We demonstrate saturation of the exciton in an encapsulated sample with single pulse reflectance, an experiment recently presented in [146]. We plot reflectance measured with beams having powers between 2 and 640 μW, a range over which the sample is not damaged. Reflectance of the sample with a beam having a power of 640 μW indicates that this power is sufficient to saturate the exciton response.

7.5 Summary

In summary, we have measured a significant improvement in sample consistency and stability by encapsulating monolayer $MoSe_2$ in hBN. We have demonstrated a homogenous linewidth narrowing by sample encapsulation that is expected for the change in the dielectric environment. This is evidence for the measured homogeneous linewidths of the encapsulated monlayer $MoSe_2$ being radiatively limited and thus

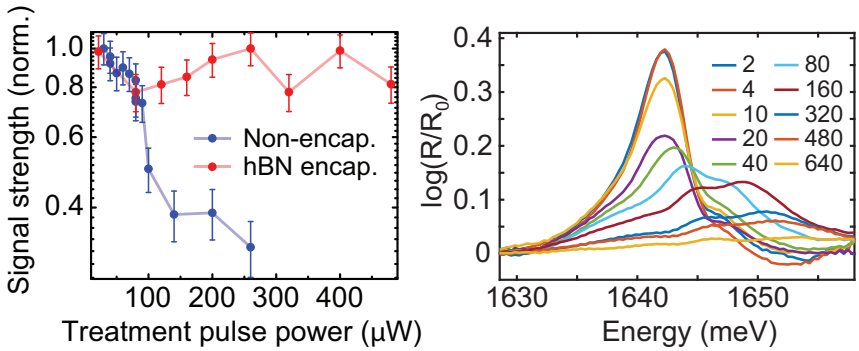


Figure 7.3: (left) Samples are exposed to a treatment beam of pulses for one minute and turned off. MDCS signal strength is measured as a function of this treatment pulse power, where the measurements are made in ascending order. We find lasting sample damage to non-encapsulated samples plotted in blue, while the encapsulated sample in red is resilient up through powers that saturate the exciton. (right) Single pulse reflectance spectroscopy is used to demonstrate saturation of the exciton in encapsulated samples at high powers that do not damage the nonlinear exciton response.

consistently lacking pure dephasing. The encapsulated monolayers are resilient to temperature cycling and high-power resonant optical excitation. Non-encapsulated monolayers on sapphire, however, are very inconsistent and susceptible to damage.

CHAPTER VIII

Treatment of Inhomogeneity for Radiative Limited Excitons in Monolayer $MoSe_2$

8.1 Introduction

In the previous chapter we discussed measuring homogeneous linewidths in monolayer $MoSe_2$ that were nearly radiatively limited at low temperatures. Measuring radiatively limited linewidths implies that the excitons in these samples are not strongly scattered to localization sites or dark states, which is very distinct from semiconductor nanostructures like GaAs quantum wells that we are used to measuring. In this chapter we identify and measure effects that arise in this regime where radiative decay dominates the sample response, and the sample emission is dominantly coherent.

8.2 Critical Coupling

The homogeneous absorption $A(\omega)$ of an optically thin single crystal is [55, 153, 51]:

$$A(\omega) = \frac{2\gamma_{sca}\gamma_{rad}}{(\omega - \omega_0)^2 + (\gamma_{sca} + \gamma_{rad})^2}, \tag{8.1}$$

where γ_{sca} is the non-radiative decay rate introduced by scattering mechanisms, γ_{rad} is the radiative decay rate, and ω_0 is the center frequency of the exciton resonance. We see that without scattering no absorption occurs, which had also been predicted

be J. J. Hopfield in his theory of excitons [49]. The absorption is maximized on resonance when $\gamma_{sca} = \gamma_{rad}$, which is called the critical coupling condition. We also see that at the critical coupling condition the absorption is only 50%. This is the peak absorption that can be realized by a suspended thin absorbing layer.

The critical coupling condition is difficult to realize in nanostructures, like quantum wells, where exciton localization introduces significant non-radiative decay mechanisms. However, critical coupling has been realized in bulk materials. In these experiments, phonons are the dominant scattering mechanisms, and their contribution can be suppressed by lowering the sample temperature to near 0 K. Peak absorption decreases with temperature below a critical temperature in cadmium selenide single crystals with thicknesses between 300 and 800 nm [154] and in gallium selenide single crystals with thicknesses between 26 and 97 μm [155]. The effect is attributed to polaritonic effects, or strong coupling between the photon and exciton. The quantity defining the coupling between the photon and exciton is the radiative linewidth. The strong-coupling regime is where this quantity dominates the scattering processes.

Having measured radiatively limited dephasing rates in encapsulated monolayer $MoSe_2$, we should now be able to realize these effects in a two-dimensional semiconductor nanostructure.

8.3 Monolayer $MoSe_2$ Device Mounted on a Mirror

In order to measure the absorption of a monolayer of hBN encapsulated $MoSe_2$, we mount the sample on a silver mirror with a dielectric spacing of $\lambda/4$. The reflected beam is equal to the sum of the transmitted and reflected beams with a relative delay of $\lambda/2$. The coherent linear response of the sample therefore cancels, and the linear reflectance measurement is a direct measure of linear absorption. Measurements of absorption typically require measurement of both transmitted and reflected signals [147]. This configuration leads to a few additional effects that must be considered

for calculations. First, the field at the monolayer is significantly enhanced relative to the sample mounted on sapphire. The increased field strength leads to a radiative linewidth that is approximately 1.8 times larger for the mirror sample than the sample mounted on sapphire. Second, the interference of the incoming and reflected beams leads to a sample absorption at the resonance that exceeds the 50% limit for a suspended monolayer.

8.4 Linewidths at the Radiative Limit

Using linear reflectance, Jason measures an integrated absorption at 270 K that is 17.1 meV·100%, and this is not the value's limit. For a homogeneous optically thin sample, integrated absorption A is directly related to the radiative ($\gamma_{rad} = \Gamma_{rad}/2$) and scattering ($\gamma_{sca}$) linewidths by:

$$A = \frac{2\pi\gamma_{rad}\gamma_{sca}}{\gamma_{rad} + \gamma_{sca}}. \tag{8.2}$$

At high temperatures, where $\gamma_{sca} \gg \gamma_{rad}$, the integrated absorption approaches a value proportional to the radiative linewidth. For a sample mounted $\lambda/4$ from a mirror, the absorption is nearly doubled. We can therefore estimate from the high temperature total absorption, A_0, that the effective radiative linewidth (which is defined as the half width at half maximum of a Lorentzian) of the whole system is $\gamma_{rad,total} \approx \frac{A_0}{2\pi} = 2.7$ meV. This approach has been used to justify that photoluminescence linewidth measurements are radiatively limited [147], but the assumed homogeneous model breaks down for an inhomogeneous distribution. Below I use this value as an effective radiative linewidth, but the sample is of course not just a suspended monolayer or homogeneous.

Using MDCS we measure the inhomogeneous and homogeneous linewidths of the sample. Using the full fit function described in Chapter II we determine the sample

has an inhomogeneous linewidth with Gaussian $\sigma = 2.76$ meV. Shown in Fig. 8.1, we measure the homogeneous linewidth extrapolated to zero temperature and zero power (γ_0) and it is 0.42 meV. Though this value is not necessarily radiatively limited, it is an upper bound on the radiative linewidth. We compare this linewidth to the

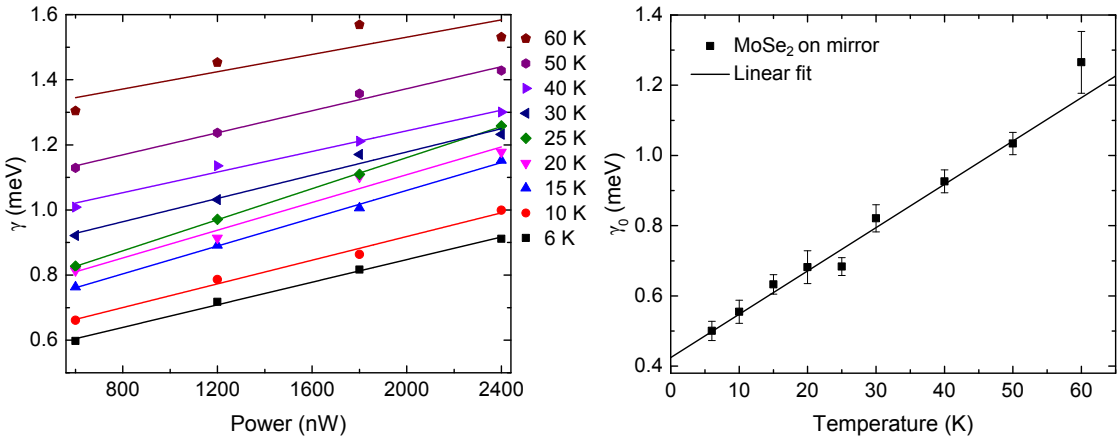


Figure 8.1: We measure MDCS of $MoSe_2$ sample mounted $\lambda/4$ from a silver mirror as a function of excitation power and sample temperature. On left: homogeneous linewidth is plotted as a function of power per excitation beam as temperatures in the range of 6 K to 60 K. We fit each power dependence to extrapolate a zero-power linewidth. On right: zero-power homoegeneous linewidth is plotted as a function of temperature. We fit this to extrapolate the zero-temperature linewidth.

linewidths of 1) encapsulated samples mounted on sapphire and 2) non-encapsulated samples. These different sample configuration are expected to have different radiative linewidths due to the decreased field strength at the sample and different dielectric environment, respectively. Comparison of encapsulated and non-encapsulated samples is described in Chapter VII to justify that the encapsulated samples and the narrowest non-encapsulated samples are radiatively limited. Here we compare samples with similar dielectric environments and different field enhancement factors. For encapsulated $MoSe_2$ on sapphire the enhancement is $|C_{\text{sapph}}|^2 = 0.84$. For the encapsulated $MoSe_2$ sample mounted $\lambda/4$ from a silvered mirror the enhancement is $|C_{\text{Ag}}|^2 = 2.75$. The extrapolated zero-temperature and zero-power linewidths of these samples are 0.23 meV and 0.42 meV, respectively. These homogeneous linewidths are related to

the radiative linewidth, enhancement factor, and scattering rate at 0 K ($\gamma_{sca,\,0K}$) by:

$$\gamma_0 = \gamma_{rad,\,C=1} \cdot C + \gamma_{sca,\,0K}. \tag{8.3}$$

Assuming the mirror does not affect the scattering rate, we determine that the ratio of the measured radiative linewidths of the mirror and sapphire mounted samples should be $C_{\mathrm{Ag}}/C_{\mathrm{sapph}} = 1.8$. This is nearly the ratio of the homogeneous linewidths that we measure using MDCS (1.83), and so we conclude that $\gamma_{sca,\,0K} = 0$. With the comparisons in Chapter VII and here we are able to convince ourselves that at the center of the inhomogeneous distribution, the homogeneous linewidth is radiatively limited.

The logical progression is to assume that the other linewidths in the inhomogeneous distribution are also radiatively limited. The linewidth would thus scale with the amplitude of the signal in the distribution, but this would also implicitly assume that the oscillator strength (μ^2) depends on energy. We hit a fork in the road because we must now consider how the measured inhomogeneous distribution scales with the oscillator strength.

If we assume the nonlinear signal scales linearly with the inhomogeneous distribution, which we typical do, then the measured $\sigma = 2.76$ meV should correspond to the linewidth of the distribution of a linear spectrum as well. The low temperature linear absorption has a linewidth $\sigma = 2.6$ meV. The linear absorption linewidth at low temperature is dominated by inhomogeneity, and our measured distributions are similar. The difference is most likely due to long-range sample inhomogeneity and our measurement of different spots on the sample.

We next perform a similar analysis with the assumption that the inhomogeneous distribution is directly related to an inhomogeneous dipole moment. Since MDCS is a four-wave mixing (FWM) spectroscopy, the amplitude of the radiated signal scales

with μ^4. This means that if we do not assume a homogeneous μ in our model, the distribution measured with FWM would be narrower than the distribution measured with a linear technique. In this case, the MDCS measurement would predict a linear absorption linewidth of $\sqrt{2} \times 2.76$ meV $= 3.90$ meV. This is much broader than the linear absorption measurement. We thus reject any theory with nonlinear scaling of the distribution, and we preserve the typical assumption made in MDCS analyses that the inhomogeneous distribution we measure should agree with distributions measured with a linear technique.

With these thoughts in mind, let's consider what a two-dimensional spectrum would look like if the radiative linewidth scales with μ^2. Since we have decided that the distribution also scales with the linear signal (also μ^2), the radiative linewidth should scale with the signal amplitude. The measured homogeneous linewidth, however, does not scale at all with the amplitude. In fact, the homogeneous linewidth is constant across the entire distribution within at least 10%. The constant linewidth would imply that the linewidths that are not at the center of the distribution are not actually radiatively limited. The difference between the homogeneous linewidth and radiative part of the linewidth at the tails is made up by scattering processes including spectral diffusion and dephasing.

However, if we were measuring distribution tails that were actually radiatively limited, we would measure a diagonal slice in MDCS that would be quite strange. Since the radiative linewidth scales with the signal strength, the amplitude should not actually change. So without broadening the tails by scattering processes, the amplitude of the signal on the diagonal of a two-dimensional plot would always be at a maximum. In Fig. 8.2 we plot the radiatively limited (in light colors) and homogeneous (in darker colors) lineshapes at various point in the inhomogeneous distribution. The linewidths of the homogeneous curves are all Lorentzian distributions with a half width at half max (HWHM) linewidth of 0.42 meV. The integral of each radiatively limited curves

is identical to the integral of the corresponding homogeneously broadened curve. The black curve follows the homogeneously broadened curves, corresponding to the measured inhomogeneous distribution.

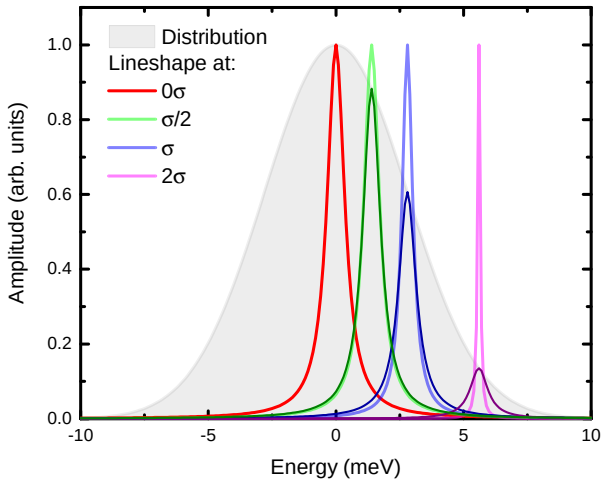


Figure 8.2: Radiatively limited lineshapes (bright colors) are plotted for various energies in an inhomogeneous distribution (black curve). Center Lorentzian feature at 0 meV has a linewidth $\gamma = 0.42$ meV and the inhomogeneous distribution has a $\sigma = 2.76$ meV. The bright curves centered at various points in the distribution, indicated in the legend, have narrower linewidths than the center because their amplitudes are lower. The dark curves correspond to lineshapes having linewidth equal to the center. The additional dephasing lowers the amplitudes of the curves on the tails of the distribution.

To confirm that the non-radiative dephasing is inhomogeneous, we quantify the total oscillator strength of the inhomogenous distribution. In the homogeneous model, the oscillator strength is proportional to the area under a Lorentzian with a linewidth equal to the radiative linewidth. In an inhomogeneous model, where a distribution of oscillators contribute to the total oscillator strength, it is essential to sum over all those oscillators. One way to determine the effective oscillator strength would be to create a set of discrete oscillators with distinct energies. This is not trivial because energies that are close together (roughly within the radiative linewidth of the sum of oscillators near that energy) cannot be treated as distinct. To determine

the integrated absorption that the MDCS linewidths would predict, we have instead related the effective oscillator strength to the area under the inhomogeneous distribution having the same amplitude as the radiatively limited oscillator at the center of the distribution. For example, at high temperatures the integrated absorption of an oscillator with a radiative linewidth of 0.42 meV, i.e. the red curve in Fig. 8.2, is 2.6 meV·100% (double for a sample on a mirror). A Gaussian with the same amplitude and a linewidth defined by $\sigma = 2.76$ meV has an area of 13.8 meV·100%. This is not consistent with the measured integrated absorption. Since the sample is $\lambda/4$ from a mirror, one would actually expect the total absorption by this sample to be double due to the coherent absorption, which would be 27.6 meV·100%. This is well over the measured 17.1 meV·100%, so somehow this consideration of inhomogeneity is not sufficient. We are currently collaborating with theorists Prof. Mackillo Kira here at the University of Michigan and Dr. Matthias Florian at the Universtät Bremen to develop a consistent interpretation that meets all the constraints imposed by linear and MDCS measurements. In the following section we apply additional constraints on the model by introducing the temperature dependence of the experimentally measured linewidth. We also show that a literature model of inhomogeneity breaks down because it convolves over the inhomogeneous distribution too early in the calculations.

8.5 Inhomogeneous Model and Dependence on Temperature

We seek a model that can consistently describe the temperature dependent linear absorption measurement using linewidth measurements from MDCS. In Fig. 8.3 we plot the measured peak absorption and integrated absorption at various sample positions for a range of temperatures from 4 K up to 270 K. The most notable features of these plots are the peak in absorption at 50 K, which we attribute to critical coupling, and that the integrated absorption does not go to 0 at low temperatures. A homogeneous model captures the peak resulting from critical coupling when the

radiative linewidth and non-radiative dephasing are equal. A homogeneous model cannot capture finite absorption at low temperature.

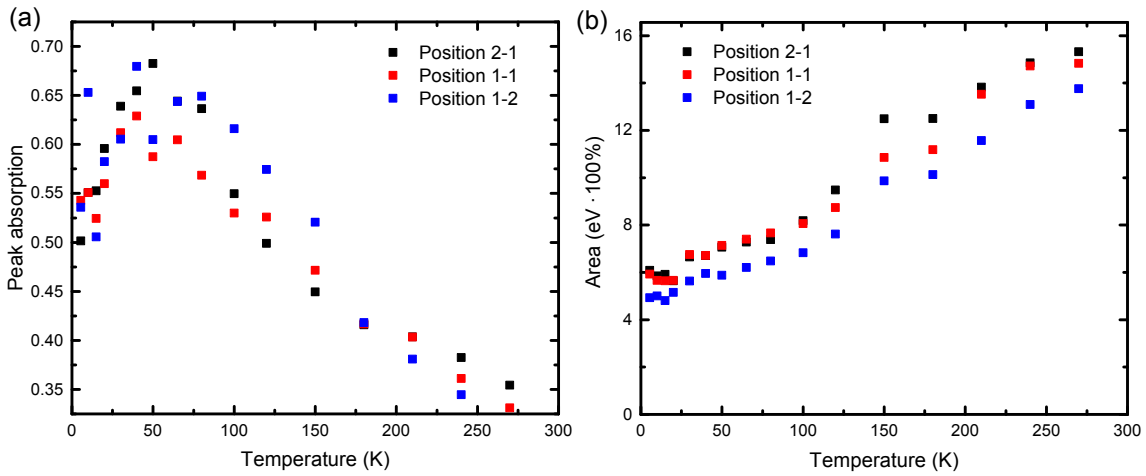


Figure 8.3: Temperature dependence of linear absorption reveals striking properties of the monolayer sample on a mirror. (a) Peak absorption as a function of temperature indicates highly coherent sample exhibiting critical coupling near 50 K. (b) Integrated absorption as a function of temperature.

MDCS measurements nearly corroborate the critical coupling determination by linear absorption, but not quite. The extrapolated 0 K homogeneous linewidth is 0.42 meV. The homogeneous linewidth increases to 0.84 meV, the point at which the radiative and non-radiative linewidth should be equal, at 35 K. This value is a little low compared to the critical coupling maximum in the linear absorption measurement. If there is any non-radiative broadening of the system initially, the peak will occur at an even lower temperature.

To attempt to model the system with inhomogeneity, we write down the reflection and transmission coefficients for a homogeneous sample:

$$r = \frac{-i\gamma_{rad}}{(\omega - \nu) + i(\gamma_{sca} + \gamma_{rad})}, \; t = 1 + r, \tag{8.4}$$

where ν is the center frequency of the resonance, defined this way for reasons that will become apparent below. We solve a simplified transfer matrix for the beam reflected

from a monolayer $MoSe_2$ sample by considering the monolayer as a two-port system, depicted in Fig. 8.4. One input port is the laser input E_0, and the other input port is the reflection off the mirror $|t|\, E_0\, e^{i\pi}$. We write down the reflection (normalized by

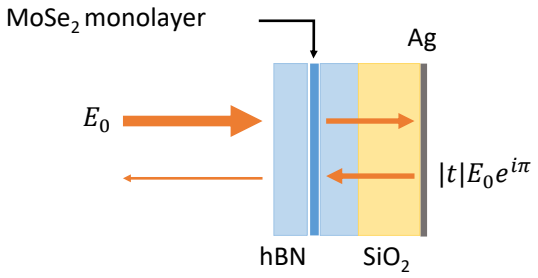


Figure 8.4: Treating the $MoSe_2$ monolayer as a two-port system we easily identify the dominant contributions to the sample response: the input beam and the transmitted beam that reflects off of the mirror with a phase delay of π. The sample is encapsulated in ∼20 nm of hexagonal boron nitride (hBN) on both sides. This heterostructure is mounted on a silver mirror coated with ∼100 nm of SiO_2. The spacing between the sample and mirror correspond to a λ/4 propagation distance for the light resonant with the $MoSe_2$ exciton.

$|E_0|^2$), which is frequency dependent:

$$R_{MoSe_2+Ag} = |r|^2 + |t|^2\,|t|^2 + 2Re\left\{(t\, r^*\, |t|)\right\}. \tag{8.5}$$

Here the first term corresponds to the light reflected from the sample, the second term corresponds to light transmitted through the sample twice due to the reflection off the mirror, and the third term corresponds to the interference of the incoming beam and beam reflected off the mirror at the sample. The last term depends on the phase of the reflected signal with respect to the incoming beam. Sample absorption for this sample is $A_{MoSe_2+Ag,\,\mathrm{hom.}} = 1 - R_{MoSe_2+Ag}$. Absorption peaks when the reflection off of the mirror is out of phase with the incoming beam. This treatment has been motivated by coherent perfect absorption studies [156], and could be equivalently done with full transfer matrix calculations.

We next treat inhomogeneity in this simulation. Inhomogeneity in semiconductors

has previously been attempted by introducing inhomogeneity into the optical susceptibility [153]. This implementation of inhomogeneity does not successfully describe spatial inhomogeneity and destroys the coherent effects that we measure. Specifically, if inhomogeneity is implemented in this way, the critical coupling effect is not realized. It is easy to understand that introducing inhomogeneity early in a calculation leads to dephasing. In the case of spatial inhomogeneity, the local response is actually homogeneous. The local response still exhibits the critical coupling effect. Inhomogeneity should be introduced in the far field, and so convolutions with the inhomogeneous distribution must be applied after calculating the homogeneous responses.

Our initial attempt at treating inhomogeneity has been to define the homogeneous sample absorption so that it depends on the energy difference between the resonance center energy (ν) and the center of the inhomogeneous distribution (ω_0). We have used the following convolution integral to consider the amplitude of the linear absorption at the center of an inhomogeneous distribution:

$$A_{\text{inh.}}(\omega = \omega_0) = \frac{1}{\sqrt{2\pi}\sigma} \int d\nu \, A_{\text{hom.}}(\omega - \nu) \exp\left[-\frac{1}{2}\left(\frac{\nu - \omega_0}{\sigma}\right)^2\right] \tag{8.6}$$

In Fig. 8.5, we plot this peak absorption as a function of temperature. The temperature dependence of γ_{sca} is $\gamma_{sca} = aT + b/(\exp(E_{LO}/k_BT) - 1)$. We determine the linear temperature dependence by fitting our data measured up to 60 K. This agrees with measurements of non-encapsulated $MoSe_2$ previously performed by Jakubczyk *et al.*, so we use their measured LO phonon activation energy and coupling strength b [157]. The parameters we use for our simulations are $a = 0.0123$ meV/K, $b = 90$ meV, and $E_{LO} = 43$ meV. This corresponds to a linewidth of 20 meV at 270 K, which is also in agreement with the linear absorption linewidth measurement.

For the other parameters in the simulation of the peak amplitude we use $\gamma_{rad} = 0.42$ meV and $\sigma = [1.0$ meV, 2.76 meV]. The narrower inhomogeneous distribution better agrees with the data, and the broader inhomogeneous distribution is shown

because it is the inhomogeneity we measure with MDCS.

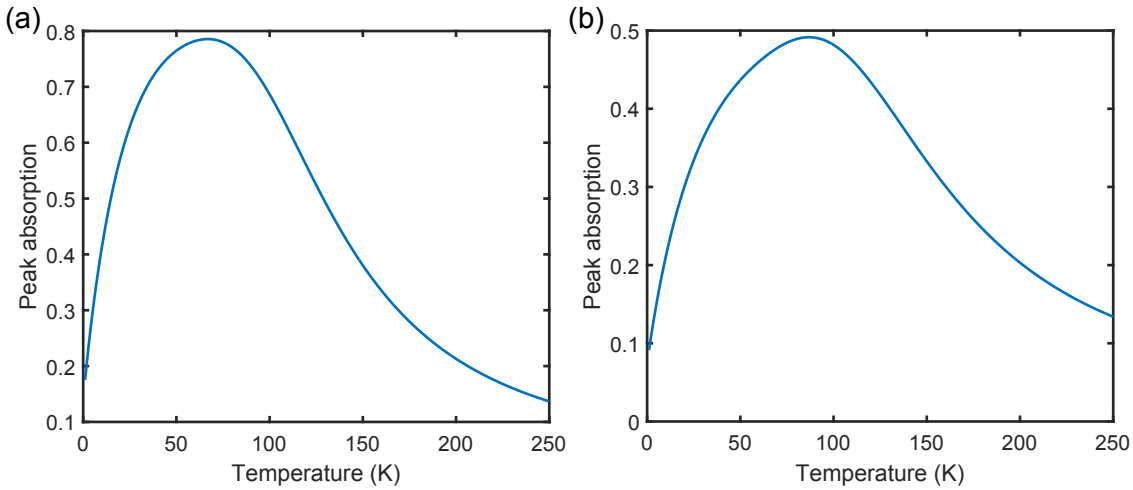


Figure 8.5: Peak absorption simulated by convolving a homogeneous response with a Gaussian to introduce the effect of inhomogeneity. Peak absorpption is plotted for (a) small inhomogeneity ($\sigma = 1.0$ meV) and (b) for inhomogeneity measured with MDCS (2.76 meV). Radiative linewidth and non-radiative dephasing used for simulation have been measured with MDCS.

The most notable finding from these simulations is that peak absorption is pushed to higher temperatures by introducing inhomogeneity, which agrees with the experimental results. We use them primarily to motivate not convolving over the inhomogeneous distribution too early in the calculation. The justification for using small inhomogeneity is perhaps due to the length scale over which inhomogeneity affects the model. This is an open question, and these simulations do not accurately predict the peak absorption at high temperatures.

Finally we attempt to model the integrated absorption. We have yet to successfully find agreement between integrals over the above equations and the measurements. The problem is that we really haven't successfully considered the effective oscillator strength realized by summing over the oscillators in the distribution. The closest model we have is to write down the integrated absorption for a homogeneous system with some effective parameters. Specifically, we use an effective radiative linewidth $\gamma_{rad,eff.} = 2.7$ meV and some initial dephasing that we attribute to inhomogeneity

(γ_σ = 1.2 meV). The use of the effective radiative linewidth is consistent with the above discussion, though a better theoretical consideration is essential to fully justify that the MDCS numbers allow for this effective radiative linewidth. The introduced dephasing due to inhomogeneity is not really justified. If the tails of the distribution are non-radiatively dephased, they could perhaps contribute to the effective initial dephasing. The model used in Fig. 8.6 is:

$$\frac{2\pi\left(\gamma_{sca}(\mathrm{T})+\gamma_\sigma\right)\times\gamma_{rad,eff.}}{\gamma_{rad,eff.}+\gamma_{sca}(\mathrm{T})+\gamma_\sigma} \tag{8.7}$$

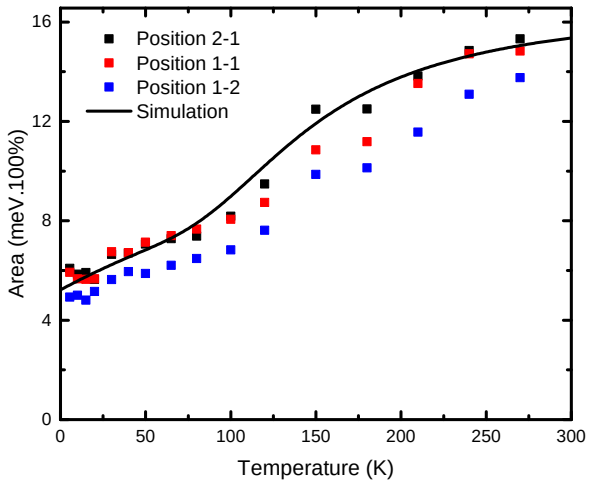


Figure 8.6: Homogeneous model with effective radiative linewidth, inhomogeneity introduced as dephasing, and measured phonon scattering with temperature.

8.6 Coherent Perfect Absorption

The absorption of the monolayer $MoSe_2$ mounted an a mirror exceeds the 50% absorption limit of a suspended monolayer. We relate this to the study of coherent perfect absorbers [158, 159, 160], which are typically implemented using metamaterials. Coherent perfect absorbers are created by counter-propagating coherent light sources so that they interfere at an optically thin material in which the dissipation is tuned so

that the material will absorb 50% of a single beam. If the counter-propagating beams interfere to form a node at the absorber, the absorber will be perfectly transparent to both beams. If the counter-propagating beams interfere to form an anti-node at the absorber, it will perfectly absorb both beams. One of the most exciting set of applications that are enabled by coherent perfect absorption is the ability to control light-with-light without nonlinearity. Using a control pulse that is coherent with a signal pulse, it has been proposed that one could restore a pulse shape, filter beam coherence, or modulate the amplitude of the signal beam [159]. The technique has also been demonstrated as a very sensitive material thickness detector [160].

One of the major challenges of coherent perfect absorbers is getting the materials thin enough. It is very difficult to make a thin absorber that will absorb 50% of a single beam. Thick materials significantly inhibit the perfect absorption bandwidth because the interference range of the counter-propagating beams is not as well defined. We therefore propose using monolayer TMDCs as a perfect absorber because it is ultrathin. Though the resonance linewidth limits the bandwidth in $MoSe_2$, the monolayer absorption we measure is unprecedented.

8.7 Summary

We have measured linear absorption and MDCS of monolayer $MoSe_2$ mounted $\lambda/4$ from a silvered mirror. This sample geometry allows us to measure the sample absorption directly by measuring the sample reflectance. We have therefore been able to quantify the frequency dependent absorption and total absorption of the sample. We realize that we are able to lower the sample temperature to below the point of critical coupling where the radiative linewidth dominates the broadening resulting from internal scattering. MDCS measurements confirm these findings, but they indicate that the sample inhomogeneity is still significant compared to the homogeneous linewidth. The quantitative linear absorption and MDCS measurements significantly constrain

a complete model to describe inhomogeneity in these monolayer semiconducting samples. We have begun to compare various models and can reject treatments of inhomogeneity that integrate over the inhomogeneous distribution to calculate the nonlinear response of the sample. MDCS measurements reveal two other aspects that constrain the model. First, the inhomogeneous distribution measured with FWM spectroscopy is identical to the distribution measured with a linear spectroscopy. Since FWM spectroscopies scale with μ^4, it is important to consider inhomogeneity in such a way that the nonlinear response does not scale nonlinearly across the distribution. Second, the homogeneous linewidth across the distribution is constant. Though the center of the distribution is radiatively limited, it is possible that non-radiative decay contributes to the linewidths of the resonances that are not at the center of the distribution.

CHAPTER IX

Coupling Quantum Confined Polaritons

9.1 Introduction

Coupling excitons in a quantum well (QW) to an optical cavity can introduce strong light-matter coupling between the two. Cavity modes and exciton resonance are strongly coupled when the band structure of the system contains anti-crossings with energy splittings that are greater than the linewidths of those resonances. In this case, the quasiparticles describing the system excitations are called polaritons. The split bands are called the upper and lower polaritons. Since polaritons have characteristics of both the cavity mode and exciton band, they have many unique properties that make them a useful system for demonstrating coherent phenomena like Bose-Einstein condensation [161].

The challenge with polaritons is increasing the interactions between polaritons. The photon-like features that give polaritons a long coherence length and a low effective mass also make polariton interactions very weak. In this chapter we describe devices that confine polaritons. By confining polaritons it is possible to significantly enhance their densities and correspondingly enhance their interactions.

MDCS is a powerful technique for measuring interactions. In this system we are able to directly measure relaxation from exciton states into polariton states. We measure self-interactions of polaritons and interaction between distinct polariton states.

We also use two-quantum MDCS to demonstrate that many-body interaction dominate the nonlinear response.

9.2 Zero-dimensional Polariton Device

Exciton polaritons are confined by a cavity structure with a finite spatial extent. The cavity is made up of 30 pairs of a distributed Bragg reflector (DBR) on the bottom and a square 7.5 μm sub-wavelength grating (SWG) on the top that is optimized as a high-reflectance mirror for light polarized along the grating bar direction. Inside the cavity, 12 GaAs quantum wells (QWs) are distributed in the three central antinodes of the cavity.

9.2.1 Sub-wavelength High Contrast Grating

High reflectors in traditional polariton devices are DBRs, which are made of alternating high and low refractive index material layers, each with an optical thickness of $\lambda/4$. In order to have a high reflectivity, which is necessary for strong coupling between the DBR-based cavity mode and the exciton, DBRs are typically made with 20-30 pairs of alternating layers.

An optical grating is a periodic structure that diffracts light and is typically used in spectroscopy as a dispersive element. For a plane wave incident on a grating with an angle θ_i, diffracted light constructively scatters from the grating with maxima at angles θ_m for integer m given by:

$$d(\sin\theta_i - \sin\theta_m) = m\lambda, \tag{9.1}$$

where d is the groove spacing and λ is the wavelength of the incident light. For $d < \lambda$, there only allowed solution is for $m = 0$ and $\sin\theta_i = -\sin\theta_0$. Thus a SWG is a reflector. The cavities here are made by a DBR stack as the bottom mirror

and a SWG as the top mirror. By careful tuning of the grating spacing and grating bar thickness, the SWGs used here have been optimized to have a broad bandwidth reflectance for light polarized along the grating that is comparable to that of a mirror made with 30 DBR pairs [162]. A high quality cavity is thus attained using a 30 DBR pairs mirror and a single SWG.

9.2.2 High Degree of Polarization

Since the grating breaks the in-plane rotational symmetry of the polariton device, SWG mirrors have high polarization selectivity. For simplicity we refer to polarization along the grating as TE polarization and across the grating as TM polarization. The gratings in these devices have been carefully optimized for a nearly perfect reflectivity for TE polarized light and low reflectivity for TM light. This means that polaritons are TE polarized and TM-polarized light excites excitons that do not strongly interact with the cavity modes. Specifically, the measured degree of polarization for lower polaritons (LPs) in these cavities is 91.9% and the degree of polarization for excitons, having a fitted orthogonal polarization to the LP, is 98.2% [163].

9.2.3 Polariton Confinement

Confinement into zero-dimensional spatial modes is enabled by the finite size of the grating. The SWG devices measured here are 7.5 μm × 7.5 μm. This finite cavity size functions as a finite square well potential that confines the polariton wave function in two directions. Though the device height and length are identical, the first vertical and horizontal excited modes are not degenerate due to the difference in the effective confinement potential, most likely resulting from the broken symmetry by the grating. This difference is evident in the spatially resolved photoluminescence (PL) measurements shown in Fig. 9.1. The first excited mode in the cross-grating direction is at a lower energy than the first excited mode in the along-grating direction.

In Fig. 9.1 we plot spatially resolved photoluminescence spectra showing discrete resonances (a) along and (b) across the SWG. Calculated energies of the LP modes agree well with measured LP resonances, but the exciton energy and effective confinement potential are both dependent on the experimental conditions. We indicate the even modes (0th, 2nd, 4th, etc.) here since the odd modes are only weakly measured for normal incidence.

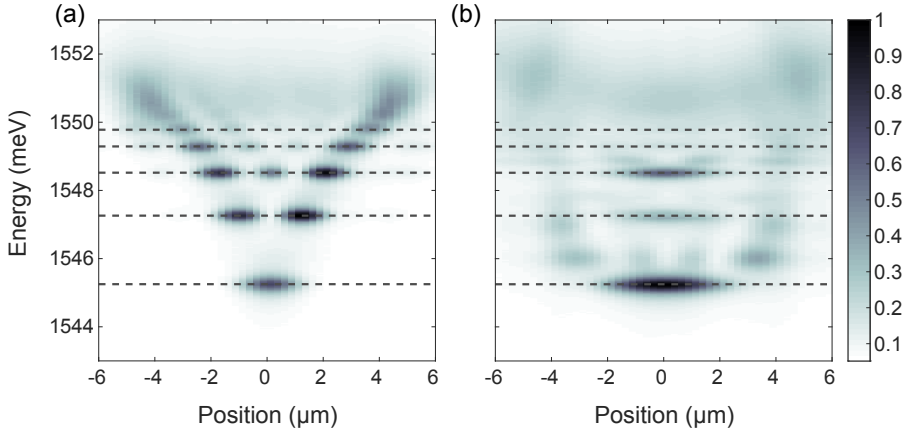


Figure 9.1: Spectrally resolved real space image of photoluminescence from the 0D cavity across (a) and along (b) the sub-wavelength grating (SWG). Dashed lines correspond to, from the bottom, 0th, 2nd, 4th, etc. modes.

9.2.4 Strong-coupling Regime and Polariton Lasing

Strong coupling between the excitons and TE polarized cavity modes has been demonstrated in these devices. Measurement of the splitting between upper and lower polariton bands reveals a strong exciton-photon coupling strength of 12 meV. As the pump power is increased above a threshold power of about 5 kW cm^{-2}, emission from the LP ground states sharply increases. Measurement of the coherence of the emission reveals that lasing has indeed been achieved in the device [163].

Since these devices confine polaritons without the use of destructive interfaces in the active media or main cavity layer, SWG polaritons enable coupling between confined states.

9.3 Measuring Coupling and Interactions

Using MDCS we measure strong polariton-polariton interactions resulting from their spatial confinement. Increasing interactions increases the rate of thermal equilibration, which facilitates polariton condensation [164, 165]. We compare double-quantum and single-quantum MDCS to demonstrate the strong self-interaction of the ground state polaritons. We also demonstrate coherent coupling between discrete polariton modes fully confined in three dimensions. Since the spatial profiles of the polariton modes are distinct, we are able to measure spatially selected coherent pathways enabling a new kind of coherent control.

Many-body effects, like excitation induced shift and dephasing, are known to be a dominant source of four-wave mixing (FWM) signals in condensed systems [7, 44, 45]. It is thus no surprise that MDCS, a set of advanced FWM techniques, offers powerful tools for directly measuring coupling and interactions. As discussed in Chapter II, single-quantum spectroscopies can reveal coupling between states. When no real states exist at a the doubled photon frequency, double-quantum signals result entirely from many-body effects. We thus use a variety of MDCS techniques to explore coupling mechanisms and their strengths in a confined polariton system. In Fig. 2.7 we illustrate one of the MDCS pulse sequences used here in which we labelled the pump as the first two excitation pulses that excite the sample into a population state and the probe as the third pulse that induces the radiated FWM signal. In this chapter we describe various ways in which we prepare the polarization or excitation position of the pump pulses differently than the probe pulse. In this way we can measure specific excitation and emission pathways in the sample.

9.4 Relaxation from Exciton States

Typical planar polariton lasers, QWs in a cavity made by top and bottom stacks of DBR layers, are optically pumped energetically above the exciton resonance and above the stop band of the DBR. In the SWG-based cavities we measure here, it is instead possible to pump the cavity on resonance with the exciton using a TM-polarized pump. The grating is nearly transparent to this polarization, and the QW absorption is very high at the exciton resonance energy. This polarization dependent front mirror reflectance is a feature for highly efficient pumping of a polariton laser, and here it enables measurement of the relaxation of exciton states into polariton states using resonant spectroscopy.

Here we compare relaxation pathways by pumping the exciton-polariton system with TE- or TM- polarized light and probing the system with TE- or TM- polarized light. The signal we measure is a third-order nonlinear response of the sample induced by three pulses, labelled A*, B, and C. Specifically this corresponds to a rephasing pulse sequence, where the coherent evolution of sample between A* and B has the opposite sign as compared to the coherent evolution of the signal emitted after C. The pumping is performed by pulses A* and B, where the exciton population created by the interference of these pulses' linear responses determines the nonlinear response. We thus spectrally resolve the absorption of the pump by measuring the phase-resolved MDCS signal as a function of the time delay (τ) between A* and B, and we Fourier transform the signal with respect to that time delay. The radiated nonlinear field is interfered with a local oscillator (LO). The LO is temporally scanned with respect to the field (time delay, t) and the interference term (the MDCS signal) is Fourier transformed as a function of the delay, t. The resulting two-dimensional plots relate the correlation between absorbed photon energies (ω_τ) and emitted photon energies (ω_t). Here we use it to measure relaxation and coupling between states that absorb light and states that emit light. For technique details see Chapter II.

In Fig. 9.2 we plot multidimensional spectra measured with a rephasing pulse sequence for various pump and probe polarizations. In Fig. 9.2(a) we plot a MDCS measurement for TE-polarized pump and probe. This corresponds to direct excitation of the polariton states and emission from those states. Some relaxation from higher energy states to the ground state is evident from this figure, and we explore this in more detail below. In (b) we measure TE pumping on the polariton states and TM probing of the exciton. Here we see that absorption of higher order polariton states modifies the exciton emission more that the lower order states. We also see that the amplitude of these features is much lower than in the other plots, and so they represent weaker effects. In (c) we measure relaxation from the exciton states pumped with TM-polarized light into the polariton states, probed with TE-polarized light. We see a dominant relaxation pathway between the exciton states and the ground state polariton. Interestingly it also seems the lowest energy exciton states significantly modify emission of highest order polariton states. This presumably results from the near degeneracy of these states. In (d) the exciton state is pumped and probed by TM-polarized light. This spectrum looks fairly different from that of a bare QW indicating that the weak coupling to the cavity is still significantly modifying the QW emission. The feature in the MDCS is strongly shifted off the diagonal, an indicator of spectral diffusion. We attribute this spectral diffusion to the weak feedback by the cavity.

The measurement of the coupling between the exciton and lowest energy polariton state is intuitively understood as relaxation. We can realize from the large vertical stripe in Fig. 9.2(c) that the absorption above the exciton is broadband and should thus have a correspondingly fast decay. By measuring the dependence of this signal on the time delay, T, between the pump pulses and the probe (not plotted), we find that the timescale of the coupling is faster than the pulse duration; that is, the formation time of the coupling peak is faster than we can measure. The spectrum decays with

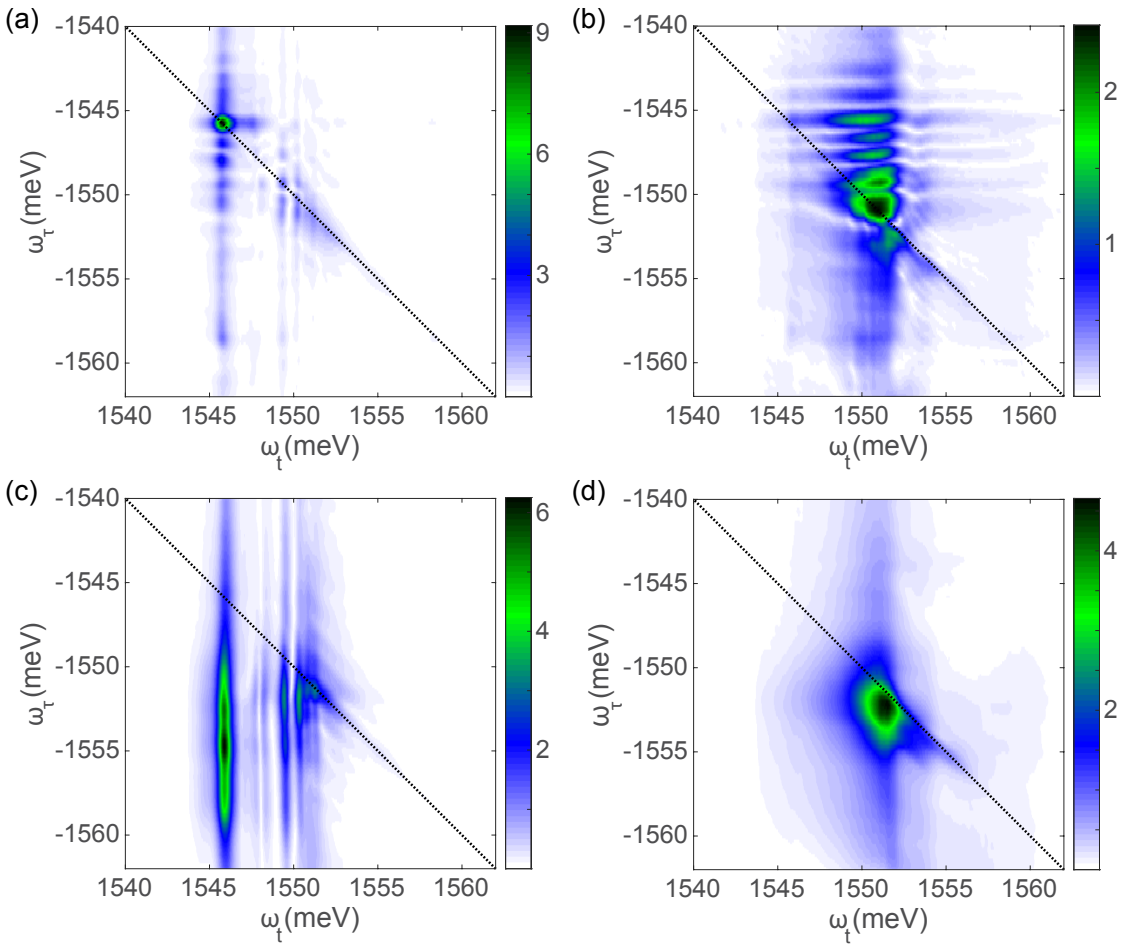


Figure 9.2: MDCS isused to measure relaxation and coupling pathways between the QW exciton and polariton states. Data are all measured with same power beams, so colorbars indicate relative signal amplitudes. The time delay between the second and third pulses, T, is 200 fs for all these measurements. (a) TE pump and TE probe correlate polariton absorption and emission. (b) With a TE pump and TM probe we measure small modification of exciton emission, and most of the modification is by the highest order polariton states. (c) With TM pump and TE probe we mostly measure relaxation from the exiton state to the ground state polariton. (d) TM pump and probe measures QW exciton weakly modified by the cavity.

the timescale of the exciton population decay.

There are, however, less intuitive phenomena revealed by these multidimensional spectra where the polariton states are nearly degenerate with the exciton. This region corresponds to the very top of the LP band where the polaritons have high momenta. In Fig. 9.2(b) we see a complete absence of polariton absorption that affects exciton

emission at the energy of the exciton resonance. We also see a pronounced cutoff of the emission at 1551 meV in (b) and (d). The cutoff implies that the polariton absorption strongly affects the emission of low energy excitons and not high energy excitons. Since the low energy excitons are the more localized excitons, this cutoff implies that the high-k polaritons couple strongly to the localized excitons. The data seem to indicate that polaritons do not couple to higher energy excitons that are more delocalized and have lower masses. Theories describing exciton-polaritons with high-mobility excitons are the current state of the art [166], and MDCS measurements could help explore these more extreme conditions.

The measurements could also be partially explained if the degree of polarization of the high-k polariton states is not as high as that of the ground state. In Fig. 9.2(d) the off-diagonal emission seems to result from two polariton resonances that we see on the diagonal in (c). This would imply that the emission of the highest polariton states is partially TM-polarized.

9.5 Polariton Interactions

In this section we explore polariton-polariton coupling with low and high spatial resolutions. For high spatial resolution scans we use an objective having a numerical aperture (NA) of 0.5 to achieve an approximately 1 um spot size. To excite all of the modes at once we perform some measurements with a lower 0.2 NA objective. For all measurements the excitation light wavevector $k = 0$, so different modes are defined spatially. To study the polaritons interaction all pulses are TE-polarized and resonant with the lowest energy polariton states.

Using the low NA objective we measure coupling between states, plotted in Fig. 9.3. In (a) and (b) we measure rephasing and non-rephasing single-quantum MDCS. For inhomogeneously broadened linewidths, like excitons, these spectra can be very different. Polaritons have very long coherence lengths, owing to the strong

coupling of the exciton to a cavity, and so the polariton linewidth is unaffected by the exciton inhomogeneity [58, 167, 168]. For this reason, rephasing and non-rephasing measurements of these confined polaritons look almost identical, even having approximately the same signal strength. What is most apparent in (a) and (b) is that all even polariton states appear to absorb light and mostly emit at the lowest energy state. Absorption by odd mode states is very low. The emission from the 2nd polariton state (along the diagonal) is also suppressed. Looking at Fig. 9.1(a) it is likely that this relatively low NA objective is still mostly collecting emission from the center of the device. In Fig. 9.1 we see PL emission from the center of the device across the grating for 0th, 4th, 8th, etc. modes. So while a large spot still excites the 2nd mode, it is understandable that the MDCS emission is not well collected and interfered with the LO.

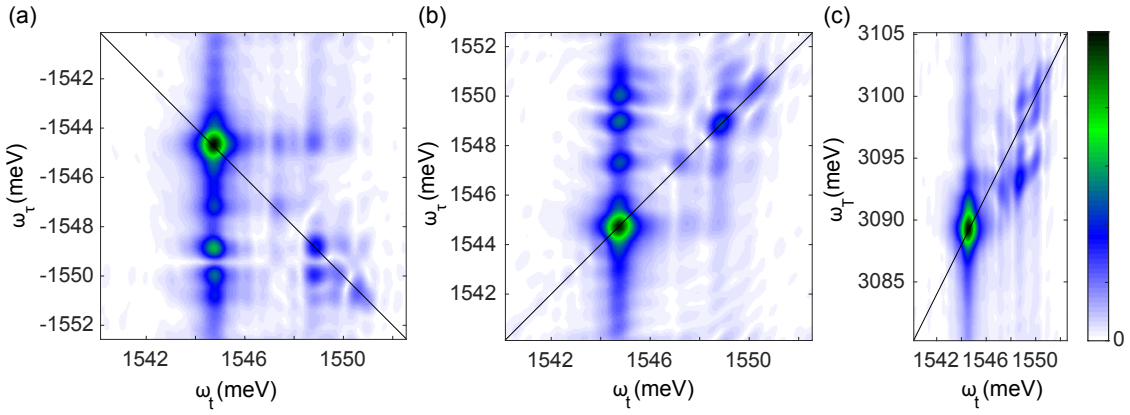


Figure 9.3: Large spot MDCS reveals absorption at many levels and relaxation into the ground state. Owing to the normal incidence of excitation we see absorption and some emission at the even modes. The odd modes are suppressed, as we would expect from the PL images. Coupling is revealed in (a) rephasing and (b) non-rephasing one-quantum MDCS. By measuring (c) two-quantum MDCS, we find the feature on the diagnoal at the 0th order state results mostly from interactions.

In Fig. 9.3 (c) we plot a double-quantum MDCS measurement of the sample measured at the same time as the singe-quantum measurements. Double-quantum spectroscopy uses a pulse sequence that correlates a single-quantum emission (or

absorption) energy to a spectrally resolved sample evolution at twice the probing frequency. For systems in which no real resonance exists at any two-photon frequency, the double-quantum pulse sequence is used to directly measure signals resulting from many-body interactions [116]. The ground state signal strength of the double-quantum measurement is comparable to that of the single-quantum measurements, which indicates that the single-quantum signal is dominated by interactions. We attribute the high polariton-polariton interaction signal to the zero-dimensional confinement of polaritons in the ground state.

From the low NA measurements, we realize strong coupling and interactions between states. These interactions seem to be strongest for the polaritons having overlapping spatial wavefunctions (e.g., the 0th and 4th states), but we also see coupling between states that do not spatially overlap. Specifically, we measure coupling between the 0th and 2nd states. Since there exists coupling between these states, we should be able to spatially select one state with the absorbing pulses and see energy transfer to another resonance at a different location. We do this by tightly focusing (NA = 0.5) the pump beams onto a different location than the probe beam. We use a spatial scanning scheme like to the one described in Chapter V in which we image a tilting mirror onto the back of the focusing objective. By scanning the beam angle at the back of the objective, we can change the focus position of one beam (i. e. the pump beam) but maintain its momentum k = 0.

In Fig. 9.4 we plot MDCS measurements for various pump and probe positions on the SWG. In (a) both the pump and probe are tightly focused onto the center of the SWG, directly exciting and measuring the ground state polariton. In (d) the pump and probe are both focused approximately 1.5 μm to the right and 0.5 μm below the center resonance. This position strongly excites the 2nd excited polariton state and does not probe the ground state. In (b) and (c) we spatially isolate coupling pathways in which we pump on one resonance and probe on the other. Specifically in (b) we

pump at the 0th polariton and measure modified emission of the 2nd polariton. In (c) we measure the inverse coupling, which is about twice as strong since relaxation to the lower energy state by phonon scattering also contributes to this signal.

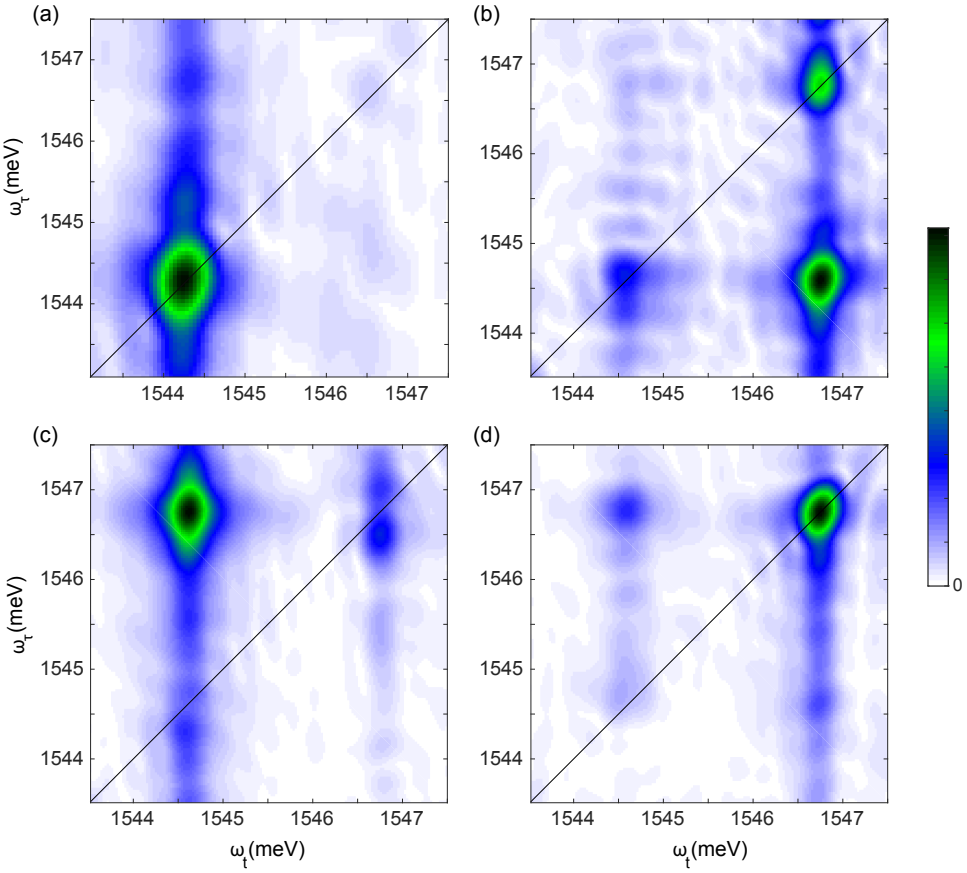


Figure 9.4: Non-local pump and probe MDCS measurements used to select coupling pathways. (a) Pump and probe excite and measure emission from the center of the SWG, corresponding to the 0th ground state polariton. (b) Pump excites ground state polariton and probe is spatially shifted relative to the pump to measure emission from the 2nd excited polariton state. Since the states are coherently coupled, the off-diagonal cross-peak between the resonances is measured. (c) Signal, dominated by relaxation, results from pump excitation of the 2nd polariton state and probing of ground state. (d) Both pump and probe spatially excite the 2nd excited polariton state and measure its nonlinear response.

Though it is interesting that we can spatially select specific coupling pathways to measure with a tightly focused beam, this result could be expected from the MDCS measurement with a large spot. We present these features to allude to a novel application: by exciting the sample with a spectrally-shaped pulse one could induce

exotic spatial emission patterns. The controlled coupling between resonance energy and position opens up the potential for coherent spatial control.

9.6 Heating Effects on Polariton Energies

In this chapter the features, and in particular the lowest energy polariton states, do not agree in every plot. When attempting to compare MDCS results to the PL, we found that the ratios of the spacing between features was maintained between plots, but the absolute spacing was not. We hypothesized that the device was modified by heating, and this modification would directly affect the polariton energies. We confirmed this result by looking at the power-dependent PL below the lasing threshold, plotted in Fig. 9.5. We found that as the devices is heated, and presumable expands, the resonances become energetically closer together. This would be expected for a decreased confinement of the wavefunction by the expansion of the device.

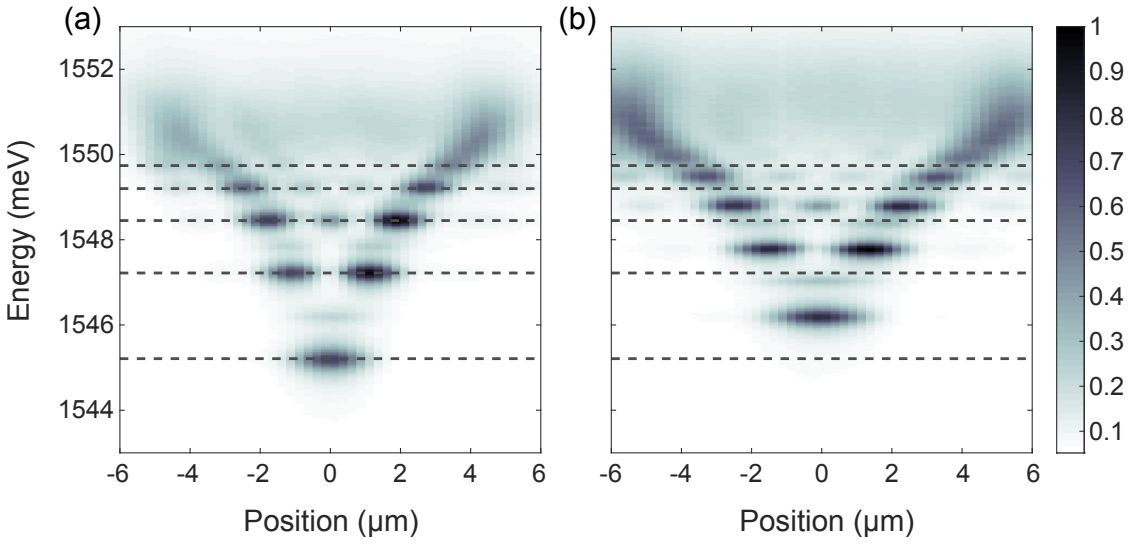


Figure 9.5: PL excited by a 740 nm pulsed laser having an 80 MHz repetition rate. (a) PL with a 2 μW excitation has a larger spacing between polariton states than (b) a 50 μW excitation. The ratio of the spacing between each resonance is the same, however.

Interestingly, the spacing between levels for the resonant MDCS measurements

was even higher than the spacing in the 2 μW excitation PL measurement. Though the MDCS excitation beam was higher, a total of 6 μW (2 μW per beam) with a repetition rate of 76 MHz, the beams seemed to heat the sample less. We attribute this decreased heating to the difference between resonant and above-gap excitation. The above-gap excitation leads to relaxation by phonon scattering, which heats the samples.

CHAPTER X

Nonlocal Four-Wave Mixing

10.1 Introduction

In nonlinear optics it is typically assumed that the response of a medium is local; that is, the induced polarization at a given point in space is only dependent on the electric field at that same point. This assumption is the basis for how we think about wave mixing, and it affects how we interpret results measured by nonlinear optical techniques. On the other hand, Wannier excitons in a 2D semiconductor are defined as eigenstates of k, and so an ideal exciton is delocalized over the crystal. We would therefore expect the nonlinear response of an ideal delocalized exciton to not be locally defined.

In the Cundiff lab we are particularly familiar with one specifically local nonlinear response: k-vector selected four-wave mixing (FWM). In the simplest realization of this class of spectroscopy, two beams having k-vectors k_1 and k_2 mix locally in a material. A population grating is formed by the interference between the two beams, and the beams can also diffract off of that population grating. To determine the effect of wave mixing by local constructive and destructive interference field interference, let's consider a third-order nonlinear response, $\chi^{(3)}$, of a material. The FWM signal corresponding to the diffraction of the k_2 beam, for degenerate beams with frequency

ω, has the form:

$$\begin{aligned} E_{FWM} &= \chi^{(3)} E_1^* e^{(-i\omega t + ik_1 x)} E_2 e^{(i\omega t - ik_2 x)} E_2 e^{(i\omega t - ik_2 x)} \\ &\quad + \chi^{(3)} E_1^* e^{(-i\omega t + ik_1 x)} E_2 e^{(i\omega t - ik_2 x)} E_2^* e^{(-i\omega t + ik_2 x)} + \text{c. c.} \\ &= \chi^{(3)} (E_1^* E_2 E_2 e^{(i\omega t - i(2k_2 - k_1)x)} + E_1^* E_2 E_2^* e^{(i\omega t - ik_1 x)}) + \text{c. c.}, \end{aligned} \tag{10.1}$$

where E fields are a function of x and t, and k and x are vectors describing a propagating fields' momenta at a given position. Assuming fields interact locally we find that the photon-echo FWM signal ($E_1^* E_2 E_2$), discussed in Chapter II, is radiated in the $2k_2 - k_1$ direction. The signal radiated in the k_1 direction is the non-rephasing FWM signal.

In FWM experiments exciton diffusion is a well known source of the breakdown of local nonlinear wave mixing [52]. This phenomenon, in which exciton population diffusion degrades the population grating created by the interference of two beams, effectively leads to wave mixing that is not local. Because of this effect on the nonlinear signal, transient grating spectroscopy can be used to measure the diffusion coefficient of excitons [169]. By measuring the population decay as a function of the angle between the exciting beams, exciton diffusion can be disentangled from population decay. However, this diffusion phenomenon is not the result of a delocalized exciton in the sense that we will refer below. Though exciton mobility and localization are related, diffusive excitons measured in transient grating experiments still have coherence lengths that are much smaller than the optically induced grating spacing.

10.2 Coherence Length and Mathematical Treatment

Every particle (or quasiparticle) has a finite spread in both position and momentum space. Since photons are massless and thus have low momenta, the minimum

coherence length of light is still quite large. This coherence length is the diffraction limit, and it is roughly half of the wavelength of light. Atoms and molecular absorption sites are typically about three orders of magnitude smaller than the wavelength of light used to excite them, so their excitations are effectively perfectly local. However, periodic structures forming bands and optical cavities can enhance the coherence length of matter. For instance excitons in ZnSe quantum wells (QWs) have been demonstrated with coherence lengths between 300 and 400 nm, which is 25-30 times larger than the exciton de Broglie wavelength [170]. Polariton condensates, in which a planar cavity enhances the spatial coherence, have been demonstrated with a coherence length that is approximately 6 μm [171].

The coherence length of a quasiparticle (i. e. an exciton) described by a band is determined by the band curvature and homogeneous linewidth. The coherence length is inversely proportional to the k-vector range the exciton subtends, and this k-vector range scales quadratically with the homogeneous linewidth, γ, and the effective mass of the exciton, m_{ex}. With appropriate factors of proportionality, the coherence length in three dimensions, l_c, is [172]:

$$l_c = \left(\frac{3\pi^2}{\sqrt{2}}\right)^{1/3} \frac{\hbar}{m_{\text{ex}}\hbar\gamma}. \tag{10.2}$$

So we see the conditions for a long coherence length: 1) a narrow homogeneous linewidth and 2) a small exciton mass corresponding to a high band curvature.

Since nonlinear wave mixing in a sample with a long coherence length is not described by traditional nonlinear optics, the mathematical treatment we employ is not conventional. There is, however, some precedent in optics for treating a beam profile as it propagates through an optical system with fractional Fourier transformss (FrFTs) [173, 174]. The FrFT is a linear transformation that generalizes the Fourier transform. It basically functions to partially transform a signal between domains. In

optics, for instance, a lens will transform a flat-top beam into a sinc function at its focus. Between the lens and the focus one could describe the beam using a FrFT with a coefficient describing the degree of the transformation that depends on the distance from the lens.

The FrFT of function f is denoted $\mathcal{F}_\alpha(k)$ and defined by:

$$\mathcal{F}_\alpha(f)(k) = \sqrt{\frac{1 - i\cot(\alpha)}{2\pi}} e^{\frac{i\cot(\alpha)k^2}{2}} \int_{-\infty}^{\infty} e^{-i\left(\csc(\alpha)kx - \frac{\cot(\alpha)}{2}x^2\right)} f(x)\,\mathrm{d}x. \tag{10.3}$$

For $\alpha = \pi/2$ this is the definition of a Fourier transform. For $\alpha = -\pi/2$ this is an inverse Fourier transform. By applying an α between 0 and $\pi/2$, we can transform a signal into a distribution that is a function of both x and k. In our treatment of nonlinear mixing with high spatial coherence, we determine the nonlinear field interaction after transforming the fields into a domain that is partially x and k. The degree of transformation is easily related to a coherence length in which the correlation of adjacent positions prevents nonlinear signals from being entirely localized.

10.3 Simulating Nonlocal Four-wave Mixing

To determine the effect of a long coherence length on FWM we explore wave mixing of fields having minimal spatial overlap. In our simulations we define a grid that is 100 times larger than the spot size (defined here as the full width at half maximum) we use to excite a nonlinear signal and with a step size that is 100 times smaller than the spot size. The finite grid allows us to use the discrete FrFT, where the order of the Fourier transform is now defined with $a = \frac{\alpha}{\pi/2}$ [175]. Specifically we use a fast implementation of this algorithm that has been written in Python [176]. Conveniently, the grid we have defined results in a coherence length equal to the spot size for $a = 0.5$. Since this parameter is dependent on the grid used to model the

system, we plot this quantity in terms of coherence length and in units of the spot size. These are absolute quantities that do not depend on simulation constraints.

Some problems can begin to arise for coherence lengths on the order of the simulation size. These are solved by defining reflecting boundaries and moderately careful selection of the excitation beam positions and momenta.

In Fig. 10.1 we plot resulting spatial domain FWM for various excitation beam positions and sample coherence lengths. In each plot the second beam E_2, which both creates a population with the first beam E_1 and induces emission of the FWM, is centered at 0. In the leftmost plot, E_1 excites the sample at the same location as E_2. Since the mechanism of signal mixing in different domains may differ, there is little reason to devote much time to considering the amplitude difference between these emission. It is, however, notable that the width of the FWM emission induced for low coherence lengths (0.1 times the excitation spot size) is narrower than the excitation spots. For large coherence length, the emission is wider. This occurs because FWM is stronger at the peak of a distribution than at the tails, and so it will narrow a linewidth in the mixing domain. For local mixing, the nonlinear signal is narrower in position. This is often touted as an advantage of multi-photon microscopy. For mixing in k, the distribution of the nonlinear signal is actually broader in position space because it is narrower in momentum space.

In the middle and rightmost plots of Fig. 10.1 we show simulated FWM for spatially separated excitation beams. In the middle plot these beams are separated by one spot size. For a low coherence length we see only FWM at the point of overlap between the beams. As the coherence length is increased, the center position of the FWM approaches the expected nonlocal FWM position of $2x_2 - x_1$. In the rightmost plot the excitation beams are two spot sizes apart, and there is hardly any spatial overlap of the beams. Using the same nonlinear coefficient as the other plots in the figure, we still find a significant nonlocal signal for a high coherence length.

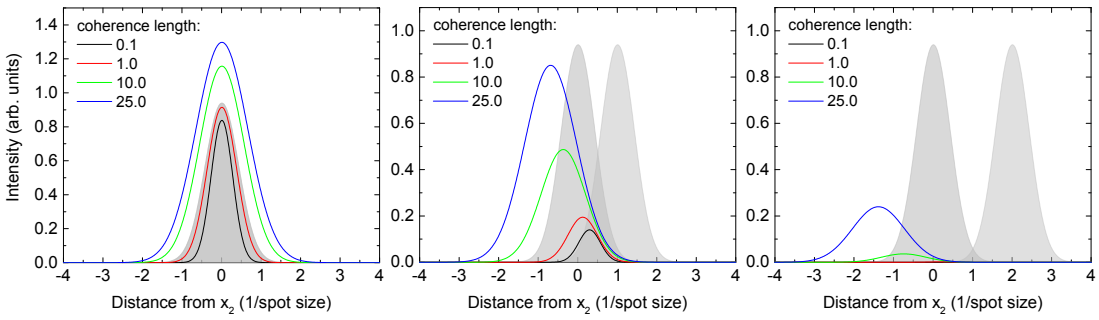


Figure 10.1: Excitation beams E_1 and E_2 excite the sample at postions x_1 and x_2. Gray shaded regions indicate excitation beams separated by 0, 1, and 2 times the spot size for the left, middle, and right plots, respectively. FWM with an arbitrary intensity is plotted for various coherence lengths where a low coherence length corresponds to local nonlinear wave mixing and a high coherence length corresponds to nonlinear mixing in momentum space.

To definitely measure a nonlocal FWM effect, it would be ideal for the FWM signal to be emitted from a position that is on the opposite side of x_2 from x_1. Any shift in the signal emission position from the point with the highest beam overlap in the sample is evidence that the nonlinear mixing is not truly local. Coherent emission that is shifted to the opposite side of the E_2 beam from the point of overlap is very strong evidence for nonlocal FWM. Such a quantity is also very easy to measure. In Fig. 10.2 we plot the simulated distance of the FWM emission from x_2 as a function of the sample coherence length. We plot this for $\Delta = x_1 - x_2$ of 1 and 2 times the spot size. We find that both relative excitation beam placements result in the same crossing point of the FWM emission and x_2. This point corresponds to a coherence length of excitons that is 2.3 times larger than the excitation spot sizes.

10.4 Experimental Result

We have measured a first test of nonlocal FWM in a confined polariton sample, the sample discussed in Chapter IX. This sample is highly nonlinear, owing to the strong confinement of the polariton modes at specific locations on the device, and

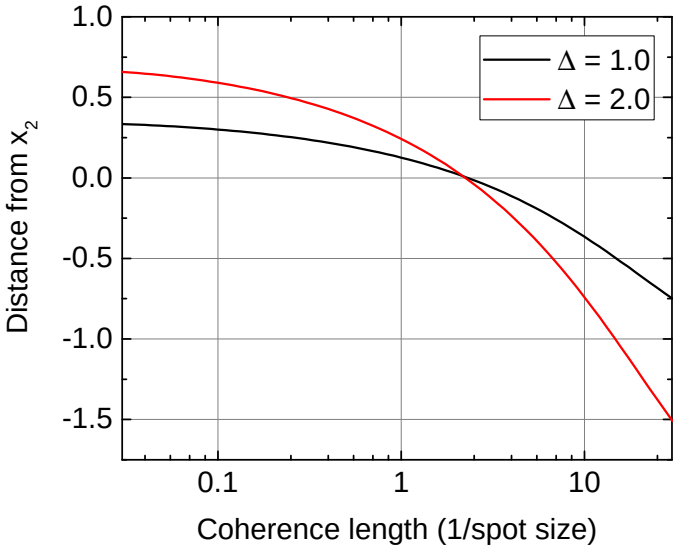


Figure 10.2: Position of FWM emission relative to the excitation beam at position x_2. The first beam is centered at x_1 and the excitation beam separation is $\Delta = x_1 - x_2$. For both beam separations the FWM emits from a position that is on the opposite side of x_2 from x_1 for sample coherence lengths greater than 2.3 times the excitation spot sizes.

these polariton modes have coherence lengths up to the size of the device size. These devices are thus a good system for demonstrating an enhancement of FWM at a $2x_2 - x_1$ position.

To measure this FWM signal we use a scheme that is similar to the experiment described in Chapter V, but we instead look at the signal corresponding to rephasing. We excite the sample with two pulsed excitation beams exciting positions x_1 and x_2. The sample image is interfered with a local oscillator (LO), which serves to only measure emission from the specific point that interferes with the LO. We lock-in detect on the beat frequency corresponding to the rephasing signal: $\omega_{FWM} = 2\omega_2 - \omega_1 - \omega_{\mathrm{LO}}$.

In a most straightforward experiment to mimic the simulations above one would fix the excitation beams and spatially scan the LO over the sample image on the detector. As the LO is scanned over the position corresponding to $2x_2 - x_1$, the nonlocal signal

would peak. In our first test, however, we have done a nearly equivalent experiment in which we instead scanned the relative postions of the excitation beams. This allows us to use our confirmed sample imaging setup for accurately determining beam positions. It also allows us to place the LO at a position that we can define by the sample image. In this test the scanning beam (having position x_1) is moved to -1.0 μm with respect to x_2 to set the LO position. The LO beam is positioned to optimize its linear interference with the beam 1 at $x_{LO} - x_2 = -1.0$ μm. The lock-in detector is then set to measure interference between the LO and rephasing signal. This interference is recorded as a function of the x_1 measured with respect to x_2. The LO delay is also scanned over the signal to measure the integrated FWM. In this experimental design the nonlocal signal should peak when $x_1 - x_2 = x_2 - x_{\mathrm{LO}}$, which corresponds to $x_1 = 1.0$ μm In Fig. 10.3 we plot the results of this measurement. Remarkably we do see the integrated FWM signal peak for $x_1 - x_2 = 1.0$ μm. We also see the signal does not peak at all where $x_1 - x_2 = -1.0$ μm, the point at which the excitation beam and LO overlap and have the maximum linear interference.

Comparison of the experimental result to the expectation for a purely nonlocal nonlinear response reveals that there are additional terms contributing to the FWM. If the signal resulted from a single state with coherence length corresponding to a semi-local nonlinear response, the response would still be spatially Gaussian and peak at a value less than $x_1 - x_2 = 1.0$ μm. We instead measure a response that is clearly not Gaussian, which indicates that we are measuring a sum of two different nonlinear signals that have different coherence lengths and most likely different spatial mode profiles. The spatial shape of the output in these confined polariton samples is also shaped by the spatial modes of the sample, which is not fully captured by the homogeneous one-dimensional model described here.

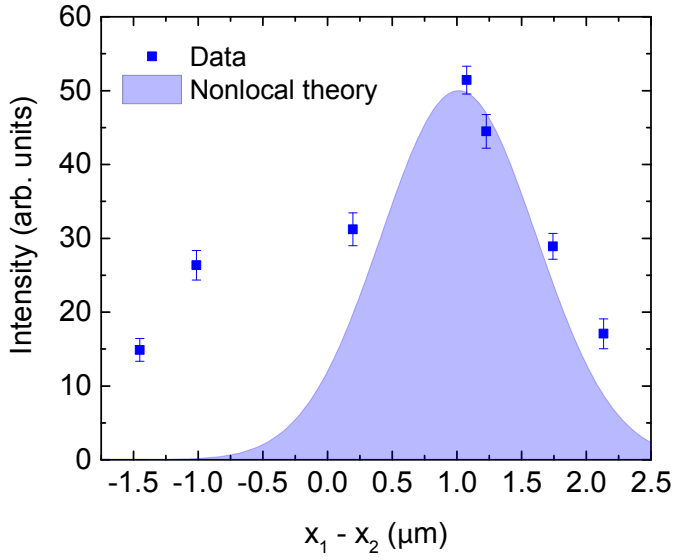


Figure 10.3: We fix a LO beam to only interfere with FWM emission from the position corresponding to $x_{LO} - x_2 = -1.0$ μm. As we scan the excitation pulse at position x_1, we find the FWM signal peak for $x_1 - x_2 = 1.0$ μm. We also plot the signal expected for nonlocal FWM that we would measure for a very high coherence length with respect to the excitation spot size. We attribute the discrepancy between the experimental measurement and nonlocal simulation to local states. We thus measure a sum of two different nonlinear responses, one of which is local and the other of which is nonlocal.

10.5 Outlook

There are some aspects of this experiment that we would like to improve to make the interpretation more straightforward and that would allow us to better identify the terms that contribute to nonlocal versus local FWM. Here we present some simple improvements that could be made to better study the confined polariton sample. We also suggest other good model systems that could also be used to study nonlocal FWM.

One issue that distinguishes these nonlocal experiments from our other heterodyne detected experiments is that here we cannot measure the phase of the FWM using the usual scheme. In the other experiments we rely on a co-propagating continuous-wave (CW) reference laser to sample path length fluctuations of the each of the

four beams necessary to measure the signal. Measurement of the relative phase of the reference laser propagating along two paths requires that the beam along both paths interfere at a detector. In a nonlocal experiment, this linear interference term will inherently be very small if the reference propagates along the same path as the excitation beams. The reference laser co-propagating beam 1 has to interfere with the one co-propagating beam 2. The beam 2 reference also has to interfere with the LO reference. None of these beam pairs is spatially overlapped in a nonlocal FWM experiment. Without referencing the intereferometers, we cannot determine the signal phase, and thus we cannot spectrally resolve the signal in the typical way. The easiest way to resolve this issue is to not spectrally resolve the beam for Fourier transforming it. Rather, the signal could be dispersed and filtered in a monochromator, and the signal could be spectrally resolved by scanning the monochromator over the various polariton resonances. This would be sufficient to determine if the local and nonlocal signal contribution arise from spectrally distinct polaritons.

It would also be interesting to determine the dependence of the local and nonlocal nonlinear responses on excitation field strength. Though the $\chi^{(3)}$ term has a cubic response to the field in both cases, high field strengths will saturate the FWM signal resulting from higher order nonlinear terms. It would be interesting to measure how the local and nonlocal field saturate, which could help shed light on the total oscillator strength of both types of excitation.

The other aspect of this experiment that hinders the straightforward interpretation has been the use of a confined polariton sample. The sample has the advantage of having a long coherence length in which there are single spatial modes with excitation spots separated by several microns. The disadvantage is that the spatial mode we excite has lobes that spatially filter the absorption and emission. In the simulations, the emitted FWM has a beam shape defined by the excitation spot and broadened by the nonlinear mixing in momentum space. An identical experiment would require

a spatially homogeneous sample that also would not spatially filter the nonlinear interacting fields. We thus suggest this experiment be performed in a planar polariton sample. Lower polaritons in planar samples have high curvature and can have long photon lifetimes up to 100-200 ps [177]. These high-quality-cavity polaritons thus have very long coherence lengths that are sufficient to measure highly nonlocal FWM.

Measuring an exciton that exibits nonlocal FWM requires is less trivial. There is currently work being done in the lab to acquire and characterize nearly perfect quantum well (QW) samples. One push has been to design a thick GaAs QW sample that is just thin enough to maintain a separation between the heavy-hole exciton and continuum states that is greater than its linewidth. A thick QW should be less sensitive to inhomogeneity and should thus have more delocalized excitons. Experimentally, by measuring a spectral dependence of the nonlocal FWM, we can isolate the signal due to more delocalized excitons by measuring higher energy excitons. These higher energy excitons in an inhomogeneous distribution are known to be more delocalized.

CHAPTER XI

Summary and Outlook

We have demonstrated several applications of heterodyne-detected nonlinear spectroscopies for studying fundamental semiconductor properties in a diffraction-limited spot. In this thesis the techniques and interpretations are detailed so they may be applied consistently to a variety of systems. Here we study many-body effects and intrinsic properties of zero- and two- dimensional semiconductor nanostructures. Most of the experiments test hypotheses that could not have been tested by other known spectroscopy methods. Experimental results are compared with existing theories and, in some cases, used to motivate new theoretical work. Coherent spectroscopy at the diffraction limit has enabled measurement of unexplored phenomena in semiconductors, and we have only begun to explore the physics that these techniques allow us to access.

One of the most pervasive topics in this thesis has been the concept of coherence. We address both temporal and spatial coherence of the excitation source and sample response. Throughout most chapters we excite resonances using a coherent light source, like a laser. The temporal coherence of the excitation is destroyed by decay and dephasing, which result from inelastic and elastic scattering processes. The spatial coherence is destroyed by inhomogeneity in the samples. We measure these coherence decays and demonstrate how they affect the macroscopic properties of the semiconductor. MDCS, in particular, is capable of measuring scattering process and

inhomogeneity with incredible sensitive. The technique has allowed us to make strong statements about the radiative linewidth, which describes the light-matter coupling strength, of monolayer $MoSe_2$. Notably we have also shown that systems with high spatial coherence, which is enhanced by strong light-matter coupling, can exhibit nonlocal FWM. Such highly coherent systems actually introduce coherent coupling between different eigenstates over a range exceeding the excitation spot size. We experimentally demonstrate these effects in confined polaritons by showing coherent coupling between states and with spatially distinct excitation beams. Within the spot size we have shown that delocalized states with high spatial coherence also enable coupling of localized states.

Since we are pushing the resolution limit of far-field optics, we explore effects that become relevant at small length scales. We have developed an experiment to measure spatial dynamics of a diffraction-limited spot of resonantly excited excitons in a well-characterized QW sample. The experiment demonstrates effects that do not occur in larger excitation spots, implying that we are approaching an intrinsic length scale of the semiconductor. Specifically, either the optical gradient or spatial confinement of the excitation spot has an effect on the semiconductor response. To describe these effects microscopically, our theory collaborators have had to develop a rigorous theoretical treatment of the inhomogeneous semiconductor Bloch equations. This project has demonstrated another case in which consideration of many-body effects is essential for understanding the macroscopic behavior of a semiconductor. It has also motivated the development of new theoretical methods that can be applied to range of spatially inhomogeneous systems and inhomogeneous excitations.

Throughout this thesis, most chapters are concluded with an outlook describing the next steps to be done in those projects. These are presented because every answer a physics experiment provides is accompanied by new questions and new engineering solutions. There a few next big steps in coherent spectroscopy of semiconductor

nanostructures that will really advance the field. One major advance is to couple MDCS with near-field techniques for improved spatial resolution. Using near-field techniques it is possible to improve the resolution of surface states by two orders of magnitude from the diffraction limit. The other major step for accessing regimes that have not been well studied is to explore many-body physics with MDCS of samples at high excitation densities. Since MDCS is a third-order nonlinear effect, measurements are typically performed in the weak-field limit where light-matter interactions can be treated perturbatively. Using a prepulse it is possible to increase the sample excitation density and maintain relatively weak probing fields [105, 97]. An incoherent preparation pulse can also Stark shift optical transitions or induce lasing. In future works, prepulse experiments will thus be used to measure many-body interactions in entirely new regimes.

APPENDICES

APPENDIX A

Increasing Quantum Well Thickness

We have measured nonlocal FWM in a polariton sample, but the long coherence length in polaritons is really introduced by the cavity. One goal we have is to create a system in which a matter state has a very long coherence length that enables measurement of a nonlocal FWM signal. An example of such a system would be a perfect QW in which the excitons are truly eigenstates of k, and so they are delocalized over the crystal. One way we can attempt to create such an ideal QW is to actually design it to be very thick. Monolayer fluctuations at at the barrier boundaries are inevitable, but we can minimize their effect on the energetic inhomogeneity. We have already demonstrated in Chapter VI that we can distinguish resonances with monolayer differences in their confinement by making the QW very thin QW. Here we calculate how thick we can make a GaAs QW with $Al_xGa_{1-x}As$ barriers..

There are a few constraints. 1) The heavy-hole must be bound by an energy that is at least greater than its linewidth at low temperatures. If the binding energy is less than its linewidth, the exciton will be rapidly dephased by continuum states, which leads to a decreased coherence length. 2) Throughout this thesis we have mostly ignored the light-hole exciton. Though bulk GaAs the heavy-hole and light-hole bands are degenerate at the Γ point, strong confinement lifts the degeneracy in

the samples we have measured. As we weaken that confinement by making a larger QW, the two bands approach degeneracy again. Overlap with the light-hole exicton can significantly broaden the homogeneous linewidth of the heavy-hole exciton.

Since a bound light-hole exciton lies either at or below the heavy-hole continuum, the energy difference between the heavy- and light-hole excitons must be greater than the sum of their linewidth. Calculating this splitting as a function of QW width requires determining several quantities. The largest source of the difference in exciton energies is the splitting between the bands. Confinement increases the band gap of GaAs, and the increase is greater for bands with lower effective masses. This is easily understand by considering the single particle wave function in a finite square well. The wave function extends into the barrier, which increases the energy levels. The band gap increases also depend on aluminum concentration (x) because the barrier height increases with x.

There are two other terms that effect the energy splitting of the exciton transition eneries, but they are about an order of magnitude smaller than the band gap difference. These are the difference in binding energies [178] and coupling of the light-hole band to the spin-orbit split-off band [179]. The explanation for the difference in binding energies is equivalent to the explanation for the difference in band gaps. Since this is a binding term, the sign is opposite. This calculation therefore decreases the splitting between heavy- and light-hole excitons, but it is much smaller than the band gap difference. The split-off band results from spin-orbit coupling and is split from the other bands by about 0.3 eV [179]. Though the splitting is quite large, the split-off band repels the light-hole band at $k = 0$ and further decreases the energy difference between the heavy- and light- holes.

These theoretical consideration were all made long ago and experimentally tested so that it is possible to calculate QW transition energies very precisely [4]. We additional compile a few experimental measurements in Fig. A.1. To interpolate

between experimental points they are fit with the function:

$$\sum_{k=0}^{3}\sum_{j=1}^{4} A_{jk}\frac{x^k}{(L_w+L_0)^j}, \tag{A.1}$$

where L_w is the well width and all other parameters are free. The fit function is only reliable where there are data points [4].

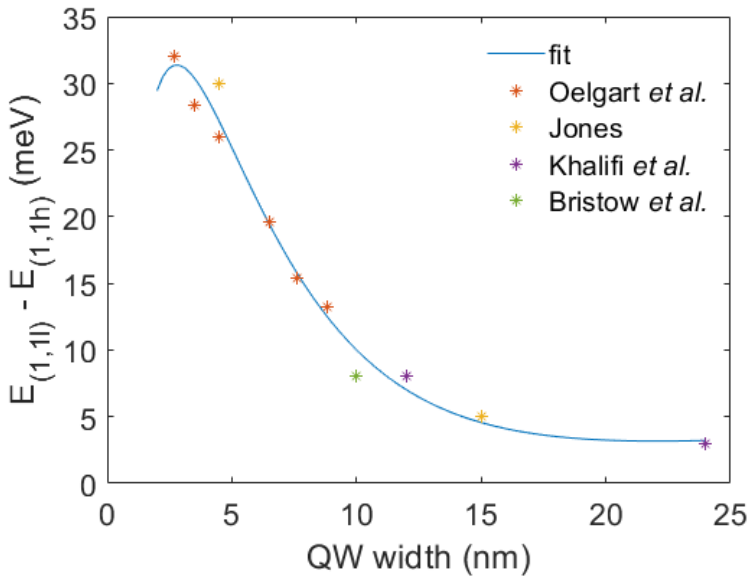


Figure A.1: We plot the measured energy splitting between the 1s excitons of the heavy-hole ($E_{(1,1h)}$) and the light-hole ($E_{(1,1l)}$) for various QW thicknesses. Results are from a variety of sources [1, 4, 5, 6]. The results are fit to interpolate the splitting energies between data points.

We have experimental precedent and theoretical confirmation that the heavy- and light-hole excitons should be split by 3 meV in a 20 nm GaAs QW with $x = 0.3$. Since we measure inhomogeneous linewidths in 10 nm QWs that are 1 meV [1], we anticipate the 3 meV splitting should be sufficient to measure a clean heavy-hole exciton. The plan for the QW samples is to measure their homogeneity, estimate the coherence length of the exciton, and then attempt to measure nonlocal FWM in the sample. Though these samples have been developed and measured before, they have not been measured with MDCS. We therefore do not know how homogeneous the

samples were. In addition, QW growth processes have improved as researchers have focused on minimizing interface fluctuations. These improvement have been largely motivated by the development and growth of quantum-cascade lasers [180, 181].

APPENDIX B

Population Rising

Interpretation of a FWM signal is not always trivial because the signal is dominated by many-body effects [7]. Rohan Singh pointed out in his thesis that the evolution of time-integrated FWM on GaAs QWs is non-monotonic [182]. Surprisingly, the FWM signal measured as a function of the time delay T does not decay exponentially and even increases on $\sim$10 ps timescales at temperatures > 50 K. Exciton spin flip and creation of excitons from electron-hole pairs in the continuum were listed as possible explanations for the increase.

To better elucidate the source of the non-monotonic dynamics, we measure spectrally resolved transient absorption of the heavy-hole and light-hole resonances on a sample containing four GaAs QWs. These data have been measured using a non-collinear transient absorption spectroscopy having a 50 μm spot size. An example spectrum is plotted in Fig. B.1 for a sample temperature of 50 K and cross-circularly polarized pump and probe pulses. The biexciton signal is negative because the pump induces absorption of the probe at these lower energies. The heavy-hole and light-hole exciton signals are positive, corresponding to saturation of those resonances by the pump pulse. Regions at the wings of the heavy-hole resonance in particular are also negative due to excitation induced broadening of the resonance.

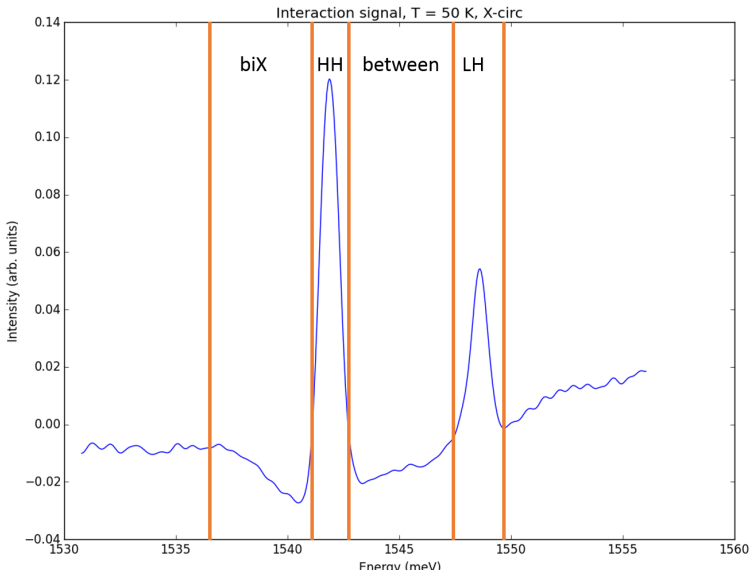
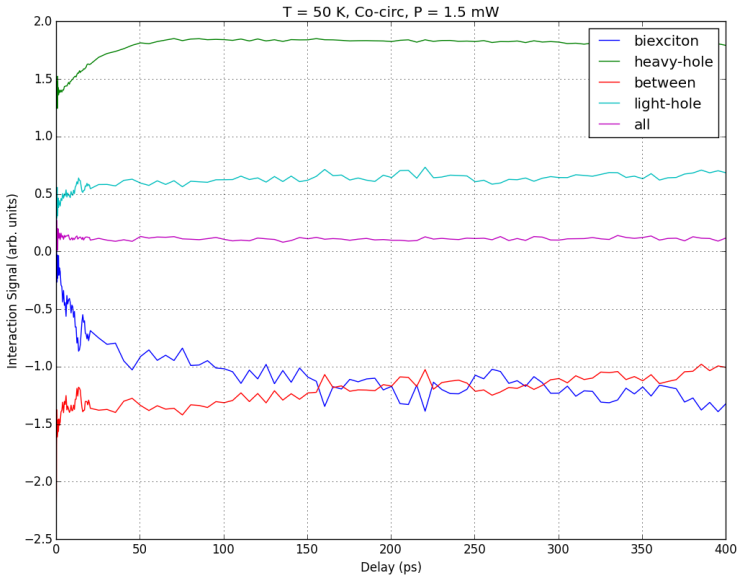


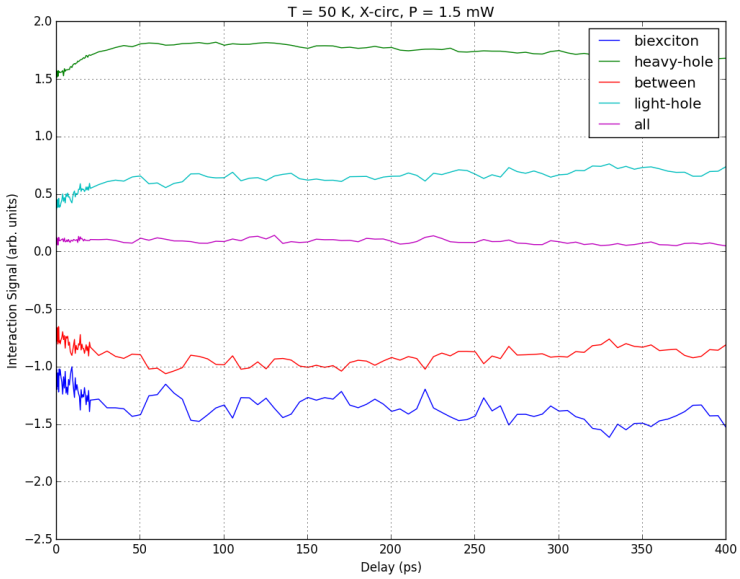
Figure B.1: Spectrally-resolved transient absorption signal measured at 50 K using cross-circularly polarized pump and probe. The signal is normalized by the laser spectrum. In the following figures we plot the integrated signal in the regions labeled biexciton (biX), heavy hole (HH), light hole (LH), and between the HH and LH.

We measure the interaction signal integrated over each spectral region and integrated over the entire spectrum. We see that the total integrated signal is fairly constant and nearly zero. This is an indicator that the signals are dominantly due to resonance shifting and broadening, not Pauli blocking. Specifically, the negative wings of the heavy-hole resonance nearly cancel the positive peak, which indicates that the area of the absorption resonance is not changed by the pump excitation. Looking at the interaction signal for the co-circularly polarized pump and probe we can determine the timescale of spin flips [183]. At time delay $T = 0$ ps, there is no biexciton excitation because all of the excitons excited by the pump and probe have the same spin. After some time, the excitons created by the pump flip their spin so that the delayed probe pulse can induce the excitation of biexcitons. We see that the signal at the heavy-hole resonance also increases with this timescale. For cross-circularly polarized pump and probe the heavy-hole resonance broadens, thereby further decreasing absorption at the exciton resonance (increases the differ-

ential absorption signal) while increasing spectral absorption of the wings.



(a)



(b)

Figure B.2: Transient absorption integrated over regions shown in Fig. B.1 of a quantum well sample at 50 K. (a) Co-circularly polarized pump and probe directly excite excitons having the same spin. They do not immediately excite biexcitons. The timescale of the biexciton signal increase is the timescale of exciton spin flipping. The exciton signals also grow with this timescale. The integrated signal over the entire spectrum is nearly constant. (b) The heavy-hole resonance grows over 30 ps while the magnitude of the resonance wings also grows. The integrated signal remains constant indicating the the heavy-hole resonance is broadening.

From these data we conclude that spectral diffusion is responsible for the non-monotonic transient absorption signals. Though the spectrally integrated transient absorption signal does not grow with time, time-integrated FWM is sensitive to spectral changes including biexciton formation and linewidth broadening. For co-circularly polarized pump and probe pulses we identify biexciton formation as the primary source of the signal rise in time delay T. We measure that biexciton formation increases the spectral absorption at the biexciton resonance and correspondingly decreases the spectral absorption at the exciton resonance. We measure that the similar effect measured with cross-circular polarized pump and probe results from broadening of the exciton resonance. The broadening effect has the same timescale as biexciton formation indicating that it too results from spin flipping of the exciton.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Alan D. Bristow, Tianhao Zhang, Mark E. Siemens, Steven T. Cundiff, and R. P. Mirin. Separating homogeneous and inhomogeneous line widths of heavy- and light-hole excitons in weakly disordered semiconductor quantum wells. *The Journal of Physical Chemistry B*, 115(18):5365–5371, 2011. PMID: 21384940.

[2] M. Kira and S.W. Koch. Many-body correlations and excitonic effects in semiconductor spectroscopy. *Progress in Quantum Electronics*, 30(5):155 – 296, 2006.

[3] Joshua D. Bell, Rebecca Conrad, and Mark E. Siemens. Analytical calculation of two-dimensional spectra. *Opt. Lett.*, 40(7):1157–1160, Apr 2015.

[4] G. Oelgart, M. Proctor, D. Martin, F. Morier-Genaud, F.-K. Reinhart, B. Orschel, L. C. Andreani, and H. Rhan. Experimental and theoretical study of excitonic transition energies in GaAs/Al_xGa_{1-x}As quantum wells. *Phys. Rev. B*, 49:10456–10465, Apr 1994.

[5] Eric D Jones. Optical determinations of energy-band dispersion curves in novel compound semiconductor materials. In *Physics and Simulation of Optoelectronic Devices III*, volume 2399, pages 515–524. International Society for Optics and Photonics, 1995.

[6] Y. El Khalifi, B. Gil, H. Mathieu, T. Fukunaga, and H. Nakashima. Dependence of the light-hole—heavy-hole splitting on layer thickness and substrate orientation in GaAs-(GaAl) As single quantum wells. *Phys. Rev. B*, 39:13533–13536, Jun 1989.

[7] D. S. Chemla and Jagdeep Shah. Many-body and correlation effects in semiconductors. *Nature*, 411(6837):549–557, May 2001.

[8] R. P. Smith, J. K. Wahlstrand, A. C. Funk, R. P. Mirin, S. T. Cundiff, J. T. Steiner, M. Schafer, M. Kira, and S. W. Koch. Extraction of many-body configurations from nonlinear absorption in semiconductor quantum wells. *Phys. Rev. Lett.*, 104:247401, Jun 2010.

[9] W.W. Chow and S.W. Koch. *Semiconductor-Laser Fundamentals: Physics of the Gain Materials*. Springer Berlin Heidelberg, 1999.

[10] Weng W. Chow and Frank Jahnke. On the physics of semiconductor quantum dots for applications in lasers and quantum optics. *Progress in Quantum Electronics*, 37(3):109 – 184, 2013.

[11] Tianhao Zhang, Irina Kuznetsova, Torsten Meier, Xiaoqin Li, Richard P. Mirin, Peter Thomas, and Steven T. Cundiff. Polarization-dependent optical 2d fourier transform spectroscopy of semiconductors. *Proceedings of the National Academy of Sciences*, 104(36):14227–14232, 2007.

[12] P. A. Franken, A. E. Hill, C. W. Peters, and G. Weinreich. Generation of optical harmonics. *Phys. Rev. Lett.*, 7:118–119, Aug 1961.

[13] Patrick F. Tekavec, Geoffrey A. Lott, and Andrew H. Marcus. Fluorescence-detected two-dimensional electronic coherence spectroscopy by acousto-optic phase modulation. *The Journal of Chemical Physics*, 127(21):214307, 2007.

[14] Stefan Mueller, Simon Draeger, Xiaonan Ma, Matthias Hensen, Tristan Kenneweg, Walter Pfeiffer, and Tobias Brixner. Fluorescence-detected two-quantum and one-quantumtwo-quantum 2d electronic spectroscopy. *The Journal of Physical Chemistry Letters*, 9(8):1964–1969, 2018. PMID: 29608071.

[15] Vivek Tiwari, Yassel Acosta-Matutes, Alastair T. Gardiner, Richard J. Cogdell, and Jennifer P. Ogilvie. Spatially-resolved fluorescence-detected two-dimensional electronic spectroscopy probes varying electronic couplings in photosynthetic bacteria, 2018.

[16] R.W. Boyd and D. Prato. *Nonlinear Optics*. Nonlinear Optics Series. Elsevier Science, 2008.

[17] M. Lindberg and S. W. Koch. Effective bloch equations for semiconductors. *Phys. Rev. B*, 38:3342–3350, Aug 1988.

[18] Ryan P. Smith. *Quantitative measurements of many-body exciton dynamics in GaAs quantum-well structures*. PhD thesis, University of Colorado Boulder, 2010.

[19] A. E. Almand-Hunter, H. Li, S. T. Cundiff, M. Mootz, M. Kira, and S. W. Koch. Quantum droplets of electrons and holes. *Nature*, 506:471, Feb 2014.

[20] K. L. Hall, G. Lenz, E. P. Ippen, and G. Raybon. Heterodyne pump–probe technique for time-domain studies of optical nonlinearities in waveguides. *Opt. Lett.*, 17(12):874–876, Jun 1992.

[21] Peter B. Fellgett. *Theory of Infra-Red Sensitivities and its Application to Investigations of Stellar Radiation in the Near Infra-Red*. PhD thesis, University of Cambridge, 1949.

[22] Tianhao Zhang, Camelia N. Borca, Xiaoqin Li, and S. T. Cundiff. Optical two-dimensional fourier transform spectroscopy with active interferometric stabilization. *Opt. Express*, 13(19):7432–7441, Sep 2005.

[23] A. D. Bristow, D. Karaiskaj, X. Dai, T. Zhang, C. Carlsson, K. R. Hagen, R. Jimenez, and S. T. Cundiff. A versatile ultrastable platform for optical multidimensional fourier-transform spectroscopy. *Review of Scientific Instruments*, 80(7):073108, 2009.

[24] Gregory D. Goodno, Gami Dadusc, and R. J. Dwayne Miller. Ultrafast heterodyne-detected transient-grating spectroscopy using diffractive optics. *J. Opt. Soc. Am. B*, 15(6):1791–1794, Jun 1998.

[25] Patrick F. Tekavec, Thomas R. Dyke, and Andrew H. Marcus. Wave packet interferometry and quantum state reconstruction by acousto-optic phase modulation. *The Journal of Chemical Physics*, 125(19):194303, 2006.

[26] Gaël Nardin, Travis M. Autry, Kevin L. Silverman, and S. T. Cundiff. Multidimensional coherent photocurrent spectroscopy of a semiconductor nanostructure. *Opt. Express*, 21(23):28617–28627, Nov 2013.

[27] Salih Salih. *Fourier Transform: Materials Analysis*. IntechOpen, 05 2012.

[28] G. Moody, M. E. Siemens, A. D. Bristow, X. Dai, A. S. Bracker, D. Gammon, and S. T. Cundiff. Exciton relaxation and coupling dynamics in a GaAs/Al_xGa_{1-x}As quantum well and quantum dot ensemble. *Phys. Rev. B*, 83:245316, Jun 2011.

[29] Rohan Singh, Marten Richter, Galan Moody, Mark E. Siemens, Hebin Li, and Steven T. Cundiff. Localization dynamics of excitons in disordered semiconductor quantum wells. *Phys. Rev. B*, 95:235307, Jun 2017.

[30] Daniel B. Turner and Keith A. Nelson. Coherent measurements of high-order electronic correlations in quantum wells. *Nature*, 466:1089, Aug 2010.

[31] Shaogang Yu, Michael Titze, Yifu Zhu, Xiaojun Liu, and Hebin Li. Observation of scalable and deterministic multi-atom dicke states in an atomic vapor, 2018.

[32] Sebastian Goetz, Donghai Li, Verena Kolb, Jens Pflaum, and Tobias Brixner. Coherent two-dimensional fluorescence micro-spectroscopy. *Opt. Express*, 26(4):3915–3925, Feb 2018.

[33] Sebastian Roeding and Tobias Brixner. Coherent two-dimensional electronic mass spectrometry. *Nature Communications*, 9(1):2519, 2018.

[34] J. Kasprzak, M. Richard, S. Kundermann, A. Baas, P. Jeambrun, J. M. J. Keeling, F. M. Marchetti, M. H. Szymanska, R. Andre, J. L. Staehli, V. Savona, P. B. Littlewood, B. Deveaud, and Le Si Dang. Bose-einstein condensation of exciton polaritons. *Nature*, 443(7110):409–414, 2006.

[35] J. Kasprzak, B. Patton, V. Savona, and W. Langbein. Coherent coupling between distant excitons revealed by two-dimensional nonlinear hyperspectral imaging. *Nat Photon*, 5(1):57–63, January 2011.

[36] E. A. Donley, T. P. Heavner, F. Levi, M. O. Tataw, and S. R. Jefferts. Double-pass acousto-optic modulator system. *Review of Scientific Instruments*, 76(6):063112, 2005.

[37] Stanford Research Systems. *MODEL SR830 DSP Lock-In Amplifier*, 2011.

[38] A. Weiner. *Ultrafast Optics*. Wiley Series in Pure and Applied Optics. Wiley, 2009.

[39] Austin P. Spencer, Hebin Li, Steven T. Cundiff, and David M. Jonas. Pulse propagation effects in optical 2d fourier-transform spectroscopy: Theory. *The Journal of Physical Chemistry A*, 119(17):3936–3960, 2015. PMID: 25880720.

[40] Patrick F. Tekavec, Jeffrey A. Myers, Kristin L. M. Lewis, Franklin D. Fuller, and Jennifer P. Ogilvie. Effects of chirp on two-dimensional fourier transform electronic spectra. *Opt. Express*, 18(11):11015–11024, May 2010.

[41] Mark E. Siemens, Galan Moody, Hebin Li, Alan D. Bristow, and Steven T. Cundiff. Resonance lineshapes in two-dimensional fourier transform spectroscopy. *Opt. Express*, 18(17):17699–17708, Aug 2010.

[42] Christopher L. Smallwood, Travis M. Autry, and Steven T. Cundiff. Analytical solutions to the finite-pulse bloch model for multidimensional coherent spectroscopy. *J. Opt. Soc. Am. B*, 34(2):419–429, Feb 2017.

[43] Amelia G. V. Spivey and Steven T. Cundiff. Inhomogeneous dephasing of heavy-hole and light-hole exciton coherences in GaAs quantum wells. *J. Opt. Soc. Am. B*, 24(3):664–670, Mar 2007.

[44] Dai-Sik Kim, Jagdeep Shah, T. C. Damen, W. Schäfer, F. Jahnke, S. Schmitt-Rink, and K. Köhler. Unusually slow temporal evolution of femtosecond four-wave-mixing signals in intrinsic GaAs quantum wells: Direct evidence for the dominance of interaction effects. *Phys. Rev. Lett.*, 69:2725–2728, Nov 1992.

[45] S. Weiss, M.-A. Mycek, J.-Y. Bigot, S. Schmitt-Rink, and D. S. Chemla. Collective effects in excitonic free induction decay: Do semiconductors and atoms emit coherent light in different ways? *Phys. Rev. Lett.*, 69:2685–2688, Nov 1992.

[46] H. Haug and S.W. Koch. *Quantum Theory of the Optical and Electronic Properties of Semiconductors*. World Scientific, 2004.

[47] C. Kittel. *Introduction to Solid State Physics*. Wiley, 1996.

[48] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. HRW international editions. Holt, Rinehart and Winston, 1976.

[49] J. J. Hopfield. Theory of the contribution of excitons to the complex dielectric constant of crystals. *Phys. Rev.*, 112:1555–1567, Dec 1958.

[50] Eiichi Hanamura. Rapid radiative decay and enhanced optical nonlinearity of excitons in a quantum well. *Phys. Rev. B*, 38(2):1228–1234, 1988.

[51] M. Kira and S. W. Koch. *Semiconductor Quantum Optics*. Cambridge Univ. Press, 1st edition, 2011.

[52] J. Hegarty, L. Goldner, and M. D. Sturge. Localized and delocalized two-dimensional excitons in GaAs-AlGaAs multiple-quantum-well structures. *Phys. Rev. B*, 30:7346–7348, Dec 1984.

[53] Yong Zhang. Motion of electrons in semiconductors under inhomogeneous strain with application to laterally confined quantum wells. *Phys. Rev. B*, 49:14352–14366, May 1994.

[54] J. Hegarty and M. D. Sturge. Studies of exciton localization in quantum-well structures by nonlinear-optical techniques. *J. Opt. Soc. Am. B*, 2(7):1143–1154, Jul 1985.

[55] Davydov A. S. and Serikov A. A. Quantum-statistical theory of light propagation in crystals. *physica status solidi (b)*, 56(1):351–363, 1973.

[56] J. P. Prineas, C. Ell, E. S. Lee, G. Khitrova, H. M. Gibbs, and S. W. Koch. Exciton-polariton eigenmodes in light-coupled $in_{0.04}ga_{0.96}As$/GaAs semiconductor multiple-quantum-well periodic structures. *Phys. Rev. B*, 61:13863–13872, May 2000.

[57] H. M. Gibbs, G. Khitrova, and S. W. Koch. Exciton-polariton light-semiconductor coupling effects. *Nature Photonics*, 5:273, Mar 2011. Review Article.

[58] D. M. Whittaker, P. Kinsler, T. A. Fisher, M. S. Skolnick, A. Armitage, A. M. Afshar, M. D. Sturge, and J. S. Roberts. Motional narrowing in semiconductor microcavities. *Phys. Rev. Lett.*, 77:4792–4795, Dec 1996.

[59] A. Vinattieri, Jagdeep Shah, T. C. Damen, D. S. Kim, L. N. Pfeiffer, M. Z. Maialle, and L. J. Sham. Exciton dynamics in GaAs quantum wells under resonant excitation. *Phys. Rev. B*, 50:10868–10879, Oct 1994.

[60] D. Gammon, E. S. Snow, B. V. Shanabrook, D. S. Katzer, and D. Park. Fine structure splitting in the optical spectra of single GaAs quantum dots. *Phys. Rev. Lett.*, 76:3005–3008, Apr 1996.

[61] D. Gammon, E. S. Snow, B. V. Shanabrook, D. S. Katzer, and D. Park. Homogeneous linewidths in the optical spectrum of a single gallium arsenide quantum dot. *Science*, 273(5271):87–90, 1996.

[62] Hervé Castella and John W. Wilkins. Splitting of the excitonic peak in quantum wells with interfacial roughness. *Phys. Rev. B*, 58:16186–16193, Dec 1998.

[63] G. Moody, M. E. Siemens, A. D. Bristow, X. Dai, D. Karaiskaj, A. S. Bracker, D. Gammon, and S. T. Cundiff. Exciton-exciton and exciton-phonon interactions in an interfacial GaAs quantum dot ensemble. *Phys. Rev. B*, 83:115324, Mar 2011.

[64] M. Kira and S. W. Koch. Quantum-optical spectroscopy of semiconductors. *Phys. Rev. A*, 73:013813, Jan 2006.

[65] S. W. Koch, M. Kira, G. Khitrova, and H. M. Gibbs. Semiconductor excitons in new light. *Nature Materials*, 5:523, Jul 2006. Review Article.

[66] M. Kira, S. W. Koch, R. P. Smith, A. E. Hunter, and S. T. Cundiff. Quantum spectroscopy with schrödinger-cat states. *Nature Physics*, 7:799, Sep 2011. Article.

[67] R. HANBURY BROWN and R. Q. TWISS. Correlation between photons in two coherent beams of light. *Nature*, 177:27, Jan 1956.

[68] H. J. Kimble, M. Dagenais, and L. Mandel. Photon antibunching in resonance fluorescence. *Phys. Rev. Lett.*, 39:691–695, Sep 1977.

[69] Y. Bromberg, Y. Lahini, E. Small, and Y. Silberberg. Hanbury brown and twiss interferometry with interacting photons. *Nature Photonics*, 4:721, Aug 2010. Article.

[70] B. L. Morgan and L. Mandel. Measurement of photon bunching in a thermal light beam. *Phys. Rev. Lett.*, 16:1012–1015, May 1966.

[71] Georgios Roumpos and Steven T. Cundiff. Multichannel homodyne detection for quantum optical tomography. *J. Opt. Soc. Am. B*, 30(5):1303–1316, May 2013.

[72] M. G. Raymer, J. Cooper, H. J. Carmichael, M. Beck, and D. T. Smithey. Ultrafast measurement of optical-field statistics by dc-balanced homodyne detection. *J. Opt. Soc. Am. B*, 12(10):1801–1812, Oct 1995.

[73] N. Peyghambarian, H. M. Gibbs, J. L. Jewell, A. Antonetti, A. Migus, D. Hulin, and A. Mysyrowicz. Blue shift of the exciton resonance due to exciton-exciton interactions in a multiple-quantum-well structure. *Phys. Rev. Lett.*, 53:2433–2436, Dec 1984.

[74] T. J. Kippenberg, R. Holzwarth, and S. A. Diddams. Microresonator-based optical frequency combs. *Science*, 332(6029):555–559, 2011.

[75] S. Chu. Nobel lecture: The manipulation of neutral particles. *Rev. Mod. Phys.*, 70:685–706, 1998.

[76] Claude N. Cohen-Tannoudji. Nobel lecture: Manipulating atoms with photons. *Rev. Mod. Phys.*, 70(3):707–719, 1998.

[77] William D. Phillips. Nobel lecture: Laser cooling and trapping of neutral atoms. *Rev. Mod. Phys.*, 70(3):721–741, 1998.

[78] David G. Grier. A revolution in optical manipulation. *Nature*, 424(6950):810–816, 2003.

[79] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and Steven Chu. Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*, 11(5):288–290, 1986.

[80] J. P. Wolfe, W. L. Hansen, E. E. Haller, R. S. Markiewicz, C. Kittel, and C. D. Jeffries. Photograph of an electron-hole drop in germanium. *Phys. Rev. Lett.*, 34(20):1292–1293, 1975.

[81] D. P. Trauernicht, A. Mysyrowicz, and J. P. Wolfe. Strain confinement and thermodynamics of free excitons in a direct-gap semiconductor. *Phys. Rev. B*, 28(6):3590–3592, 1983.

[82] V. Negoita, D. W. Snoke, and K. Eberl. Stretching quantum wells: A method for trapping free carriers in GaAs heterostructures. *Appl. Phys. Lett.*, 75(14):2059–2061, 1999.

[83] P. C. M. Christianen, F. Piazza, J. G. S. Lok, J. C. Maan, and W. van der Vleuten. Magnetic trap for excitons. *Physica B*, 249251:624–627, 1998.

[84] A. T. Hammack, N. A. Gippius, Sen Yang, G. O. Andreev, L. V. Butov, M. Hanson, and A. C. Gossard. Excitons in electrostatic traps. *J. Appl. Phys.*, 99(6):066104, 2006.

[85] T. Huber, A. Zrenner, W. Wegscheider, and M. Bichler. Electrostatic exciton traps. *Phys. Status Solidi A*, 166(1):R5–R6, 1998.

[86] A. T. Hammack, M. Griswold, L. V. Butov, L. E. Smallwood, A. L. Ivanov, and A. C. Gossard. Trapping of cold excitons in quantum well structures with laser light. *Phys. Rev. Lett.*, 96(22):227402, 2006.

[87] M. Lindberg and R. Binder. Transversal light forces in semiconductors. *J. Phys.: Condens. Matter*, 15(7):1119, 2003.

[88] R. Binder and M. Lindberg. Optical electronhole tweezers in semiconductors. *J. Phys.: Condens. Matter*, 18(2):729, 2006.

[89] K. Henneberger and H. Haug. Nonlinear optics and transport in laser-excited semiconductors. *Phys. Rev. B*, 38(14):9759–9770, 1988.

[90] A. Ashkin. Acceleration and trapping of particles by radiation pressure. *Phys. Rev. Lett.*, 24:156–159, Jan 1970.

[91] Wodzimierz Nakwaski. Effective masses of electrons and heavy holes in GaAs, InAs, AlAs and their ternary compounds. *Physica B: Condensed Matter*, 210(1):1 – 25, 1995.

[92] R. Binder, S. W. Koch, M. Lindberg, N. Peyghambarian, and W. Schäfer. Ultrafast adiabatic following in semiconductors. *Phys. Rev. Lett.*, 65:899–902, Aug 1990.

[93] S. T. Cundiff, A. Knorr, J. Feldmann, S. W. Koch, E. O. Gbel, and H. Nickel. Rabi flopping in semiconductors. *Phys. Rev. Lett.*, 73(8):1178–1181, 1994.

[94] P. Hamm and M. Zanni. *Concepts and Methods of 2D Infrared Spectroscopy*. Cambridge University Press, 2011.

[95] G. von Plessen and J. Feldmann. Temporally and spatially resolved optical experiments on semiconductor quantum wells and quantum dots. *Braz. J. Phys.*, 27/A(4):20–28, 1997.

[96] L. V. Butov, C. W. Lai, A. L. Ivanov, A. C. Gossard, and D. S. Chemla. Towards bose-einstein condensation of excitons in potential traps. *Nature*, 417(6884):47–52, 2002.

[97] Eric W. Martin and Steven T. Cundiff. Inducing coherent quantum dot interactions. *Phys. Rev. B*, 97:081301, Feb 2018.

[98] N. H. Bonadeo, Gang Chen, D. Gammon, D. S. Katzer, D. Park, and D. G. Steel. Nonlinear nano-optics: Probing one exciton at a time. *Phys. Rev. Lett.*, 81:2759–2762, Sep 1998.

[99] Xudong Fan, T. Takagahara, J.E. Cunningham, and Hailin Wang. Pure dephasing induced by excitonphonon interactions in narrow GaAs quantum wells. *Solid State Communications*, 108(11):857 – 861, 1998.

[100] D. Bimberg, N. Kirstaedter, N. N. Ledentsov, Z. I. Alferov, P. S. Kop'ev, and V. M. Ustinov. InGaAs-GaAs quantum-dot lasers. *IEEE Journal of Selected Topics in Quantum Electronics*, 3(2):196–205, Apr 1997.

[101] David Press, Thaddeus D. Ladd, Bingyang Zhang, and Yoshihisa Yamamoto. Complete quantum control of a single quantum dot spin using ultrafast optical pulses. *Nature*, 456(7219):218–221, Nov 2008.

[102] J. Berezovsky, M. H. Mikkelsen, N. G. Stoltz, L. A. Coldren, and D. D. Awschalom. Picosecond coherent optical manipulation of a single electron spin in a quantum dot. *Science*, 320(5874):349–352, 2008.

[103] Yanwen Wu, Xiaoqin Li, L. M. Duan, D. G. Steel, and D. Gammon. Density matrix tomography through sequential coherent optical rotations of an exciton qubit in a single quantum dot. *Phys. Rev. Lett.*, 96:087402, Feb 2006.

[104] Xiaoqin Li, Yanwen Wu, Duncan Steel, D. Gammon, T. H. Stievater, D. S. Katzer, D. Park, C. Piermarocchi, and L. J. Sham. An all-optical quantum gate in a semiconductor quantum dot. *Science*, 301(5634):809–811, 2003.

[105] Takeshi Suzuki, Rohan Singh, Manfred Bayer, Arne Ludwig, Andreas D. Wieck, and Steven T. Cundiff. Coherent control of the exciton-biexciton system in an inas self-assembled quantum dot ensemble. *Phys. Rev. Lett.*, 117:157402, Oct 2016.

[106] Danny Kim, Samuel G. Carter, Alex Greilich, Allan S. Bracker, and Daniel Gammon. Ultrafast optical control of entanglement between two quantum-dot spins. *Nat Phys*, 7(3):223–229, Mar 2011.

[107] S. Spatzek, A. Greilich, Sophia E. Economou, S. Varwig, A. Schwan, D. R. Yakovlev, D. Reuter, A. D. Wieck, T. L. Reinecke, and M. Bayer. Optical control of coherent interactions between electron spins in InGaAs quantum dots. *Phys. Rev. Lett.*, 107:137402, Sep 2011.

[108] Maksym Serbyn, Z. Papić, and Dmitry A. Abanin. Local conservation laws and the structure of the many-body localized states. *Phys. Rev. Lett.*, 111:127201, Sep 2013.

[109] H. C. Schneider, W. W. Chow, and S. W. Koch. Excitation-induced dephasing in semiconductor quantum dots. *Phys. Rev. B*, 70:235308, Dec 2004.

[110] S. T. Cundiff and S. Mukamel. Optical multidimensional coherent spectroscopy. *Phys. Today*, 66:44, 2013.

[111] H. F. Hess, E. Betzig, T. D. Harris, L. N. Pfeiffer, and K. W. West. Near-field spectroscopy of the quantum constituents of a luminescent system. *Science*, 264(5166):1740–1745, 1994.

[112] Valentin Delmonte, Judith F. Specht, Tomasz Jakubczyk, Sven Höfling, Martin Kamp, Christian Schneider, Wolfgang Langbein, Gilles Nogues, Marten Richter, and Jacek Kasprzak. Coherent coupling of individual quantum dots measured with phase-referenced two-dimensional spectroscopy: Photon echo versus double quantum coherence. *Phys. Rev. B*, 96:041124, Jul 2017.

[113] G. Moody, R. Singh, H. Li, I. A. Akimov, M. Bayer, D. Reuter, A. D. Wieck, A. S. Bracker, D. Gammon, and S. T. Cundiff. Influence of confinement on biexciton binding in semiconductor quantum dot ensembles measured with two-dimensional spectroscopy. *Phys. Rev. B*, 87:041304, Jan 2013.

[114] Jacek Kasprzak and Wolfgang Langbein. Coherent response of individual weakly confined exciton and biexciton systems. *J. Opt. Soc. Am. B*, 29(7):1766–1771, Jul 2012.

[115] Denis Karaiskaj, Alan D. Bristow, Lijun Yang, Xingcan Dai, Richard P. Mirin, Shaul Mukamel, and Steven T. Cundiff. Two-quantum many-body coherences in two-dimensional fourier-transform spectra of exciton resonances in semiconductor quantum wells. *Phys. Rev. Lett.*, 104:117401, Mar 2010.

[116] Lijun Yang and Shaul Mukamel. Two-dimensional correlation spectroscopy of two-exciton resonances in semiconductor quantum wells. *Phys. Rev. Lett.*, 100:057402, Feb 2008.

[117] Ajit Srivastava, Meinrad Sidler, Adrien V Allain, Dominik S Lembke, Andras Kis, and A Imamoğlu. Optically active quantum dots in monolayer WSe_2. *Nature nanotechnology*, 10(6):491–496, 2015.

[118] M Koperski, K Nogajewski, A Arora, V Cherkez, P Mallet, J-Y Veuillen, J Marcus, P Kossacki, and M Potemski. Single photon emitters in exfoliated WSe_2 structures. *Nature nanotechnology*, 10(6):503–506, 2015.

[119] Yu-Ming He, Genevieve Clark, John R. Schaibley, Yu He, Ming-Cheng Chen, Yu-Jia Wei, Xing Ding, Qiang Zhang, Wang Yao, Xiaodong Xu, Chao-Yang Lu, and Jian-Wei Pan. Single quantum emitters in monolayer semiconductors. *Nature Nanotechnology*, 10:497, May 2015.

[120] Chitraleema Chakraborty, Laura Kinnischtzke, Kenneth M. Goodfellow, Ryan Beams, and A. Nick Vamivakas. Voltage-controlled quantum light from an atomically thin semiconductor. *Nature Nanotechnology*, 10:507, May 2015.

[121] Philipp Tonndorf, Robert Schmidt, Robert Schneider, Johannes Kern, Michele Buscema, Gary A. Steele, Andres Castellanos-Gomez, Herre S. J. van der Zant, Steffen Michaelis de Vasconcellos, and Rudolf Bratschitsch. Single-photon emission from localized excitons in an atomically thin semiconductor. *Optica*, 2(4):347–352, Apr 2015.

[122] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis. Single-layer MoS_2 transistors. *Nature Nanotechnology*, 6:147, Jan 2011.

[123] Britton W. H. Baugher, Hugh O. H. Churchill, Yafang Yang, and Pablo Jarillo-Herrero. Intrinsic electronic transport properties of high-quality monolayer and bilayer MoS_2. *Nano Letters*, 13(9):4212–4216, 2013. PMID: 23930826.

[124] Yijin Zhang, Jianting Ye, Yusuke Matsuhashi, and Yoshihiro Iwasa. Ambipolar MoS_2 thin flake transistors. *Nano Letters*, 12(3):1136–1140, 2012. PMID: 22276648.

[125] K. S. Novoselov, A. Mishchenko, A. Carvalho, and A. H. Castro Neto. 2d materials and van der waals heterostructures. *Science*, 353(6298), 2016.

[126] Yu Ye, Zi Jing Wong, Xiufang Lu, Xingjie Ni, Hanyu Zhu, Xianhui Chen, Yuan Wang, and Xiang Zhang. Monolayer excitonic laser. *Nature Photonics*, 9:733, Oct 2015.

[127] Sanfeng Wu, Sonia Buckley, John R. Schaibley, Liefeng Feng, Jiaqiang Yan, David G. Mandrus, Fariba Hatami, Wang Yao, Jelena Vuckovic, Arka Majumdar, and Xiaodong Xu. Monolayer semiconductor nanocavity lasers with ultralow thresholds. *Nature*, 520:69, Mar 2015.

[128] Omid Salehzadeh, Mehrdad Djavid, Nhung Hong Tran, Ishiang Shih, and Zetian Mi. Optically pumped two-dimensional MoS_2 lasers operating at room-temperature. *Nano Letters*, 15(8):5302–5306, 2015.

[129] Andreas Pospischil, Marco M. Furchi, and Thomas Mueller. Solar-energy conversion and light emission in an atomic monolayer p-n diode. *Nature Nanotechnology*, 9:257, Mar 2014.

[130] C. R. Dean, A. F. Young, I. Meric, C. Lee, L. Wang, S. Sorgenfrei, K. Watanabe, T. Taniguchi, P. Kim, K. L. Shepard, and J. Hone. Boron nitride substrates for high-quality graphene electronics. *Nature Nanotechnology*, 5:722, Aug 2010.

[131] Xu Cui, Gwan-Hyoung Lee, Young Duck Kim, Ghidewon Arefe, Pinshane Y. Huang, Chul-Ho Lee, Daniel A. Chenet, Xian Zhang, Lei Wang, Fan Ye, Filippo Pizzocchero, Bjarke S. Jessen, Kenji Watanabe, Takashi Taniguchi, David A. Muller, Tony Low, Philip Kim, and James Hone. Multi-terminal transport measurements of MoS_2 using a van der waals heterostructure device platform. *Nature Nanotechnology*, 10:534, Apr 2015.

[132] Seongjoon Ahn, Gwangwoo Kim, Pramoda K. Nayak, Seong In Yoon, Hyunseob Lim, Hyun-Joon Shin, and Hyeon Suk Shin. Prevention of transition metal dichalcogenide photodegradation by encapsulation with h-bn layers. *ACS Nano*, 10(9):8973–8979, 2016.

[133] F. Cadiz, E. Courtade, C. Robert, G. Wang, Y. Shen, H. Cai, T. Taniguchi, K. Watanabe, H. Carrere, D. Lagarde, M. Manca, T. Amand, P. Renucci, S. Tongay, X. Marie, and B. Urbaszek. Excitonic linewidth approaching the homogeneous limit in MoS_2-based van der waals heterostructures. *Phys. Rev. X*, 7:021026, May 2017.

[134] Obafunso A Ajayi, Jenny V Ardelean, Gabriella D Shepard, Jue Wang, Abhinandan Antony, Takeshi Taniguchi, Kenji Watanabe, Tony F Heinz, Stefan Strauf, X-Y Zhu, and James C Hone. Approaching the intrinsic photoluminescence linewidth in transition metal dichalcogenide monolayers. *2D Materials*, 4(3):031011, 2017.

[135] Jakob Wierzbowski, Julian Klein, Florian Sigger, Christian Straubinger, Malte Kremser, Takashi Taniguchi, Kenji Watanabe, Ursula Wurstbauer, Alexander W. Holleitner, Michael Kaniber, Kai Müller, and Jonathan J. Finley. Di-

rect exciton emission from atomically thin transition metal dichalcogenide heterostructures near the lifetime limit. *Scientific Reports*, 7(1):12383, 2017.

[136] Galan Moody, Chandriker Kavir Dass, Kai Hao, Chang-Hsiao Chen, Lain-Jong Li, Akshay Singh, Kha Tran, Genevieve Clark, Xiaodong Xu, Gunnar Berghäuser, Ermin Malic, Andreas Knorr, and Xiaoqin Li. Intrinsic homogeneous linewidth and broadening mechanisms of excitons in monolayer transition metal dichalcogenides. *Nature Communications*, 6:8315, Sep 2015.

[137] Tomasz Jakubczyk, Valentin Delmonte, Maciej Koperski, Karol Nogajewski, Clment Faugeras, Wolfgang Langbein, Marek Potemski, and Jacek Kasprzak. Radiatively limited dephasing and exciton dynamics in mose2 monolayers revealed with four-wave mixing microscopy. *Nano Letters*, 16(9):5333–5339, 2016.

[138] Tomasz Jakubczyk, Karol Nogajewski, Maciej R Molas, Miroslav Bartos, Wolfgang Langbein, Marek Potemski, and Jacek Kasprzak. Impact of environment on dynamics of exciton complexes in a WS_2 monolayer. *2D Materials*, 5(3):031007, 2018.

[139] Kai Hao, Judith F. Specht, Philipp Nagler, Lixiang Xu, Kha Tran, Akshay Singh, Chandriker Kavir Dass, Christian Schüller, Tobias Korn, Marten Richter, Andreas Knorr, Xiaoqin Li, and Galan Moody. Neutral and charged inter-valley biexcitons in monolayer $MoSe_2$. *Nature Communications*, 8:15552, Jun 2017.

[140] Wonbong Choi, Nitin Choudhary, Gang Hee Han, Juhong Park, Deji Akinwande, and Young Hee Lee. Recent development of two-dimensional transition metal dichalcogenides and their applications. *Materials Today*, 20(3):116 – 130, 2017.

[141] Ke Chen, Rudresh Ghosh, Xianghai Meng, Anupam Roy, Joon-Seok Kim, Feng He, Sarah C. Mason, Xiaochuan Xu, Jung-Fu Lin, Deji Akinwande, Sanjay K. Banerjee, and Yaguo Wang. Experimental evidence of exciton capture by mid-gap defects in cvd grown monolayer $MoSe_2$. *npj 2D Materials and Applications*, 1(1):15, 2017.

[142] G. Moody, C. McDonald, A. Feldman, T. Harvey, R. P. Mirin, and K. L. Silverman. Electronic enhancement of the exciton coherence time in charged quantum dots. *Phys. Rev. Lett.*, 116:037402, Jan 2016.

[143] Xiao-Xiao Zhang, Yumeng You, Shu Yang Frank Zhao, and Tony F. Heinz. Experimental evidence for dark excitons in monolayer WSe_2. *Phys. Rev. Lett.*, 115:257403, Dec 2015.

[144] Ashish Arora, Maciej Koperski, Karol Nogajewski, Jacques Marcus, Clement Faugeras, and Marek Potemski. Excitonic resonances in thin films of WSe_2: from monolayer to bulk material. *Nanoscale*, 7:10421–10429, 2015.

[145] Malte Selig, Gunnar Berghäuser, Archana Raja, Philipp Nagler, Christian Schüller, Tony F. Heinz, Tobias Korn, Alexey Chernikov, Ermin Malic, and Andreas Knorr. Excitonic linewidth and coherence lifetime in monolayer transition metal dichalcogenides. *Nature Communications*, 7:13279, Nov 2016.

[146] Giovanni Scuri, You Zhou, Alexander A. High, Dominik S. Wild, Chi Shu, Kristiaan De Greve, Luis A. Jauregui, Takashi Taniguchi, Kenji Watanabe, Philip Kim, Mikhail D. Lukin, and Hongkun Park. Large excitonic reflectivity of monolayer $MoSe_2$ encapsulated in hexagonal boron nitride. *Phys. Rev. Lett.*, 120:037402, Jan 2018.

[147] Patrick Back, Sina Zeytinoglu, Aroosa Ijaz, Martin Kroner, and Atac Imamoğlu. Realization of an electrically tunable narrow-bandwidth atomically thin mirror using monolayer $MoSe_2$. *Phys. Rev. Lett.*, 120:037401, Jan 2018.

[148] F. Cadiz, E. Courtade, C. Robert, G. Wang, Y. Shen, H. Cai, T. Taniguchi, K. Watanabe, H. Carrere, D. Lagarde, M. Manca, T. Amand, P. Renucci, S. Tongay, X. Marie, and B. Urbaszek. Excitonic linewidth approaching the homogeneous limit in MoS_2-based van der waals heterostructures. *Phys. Rev. X*, 7:021026, May 2017.

[149] Andreas V. Stier, Nathan P. Wilson, Genevieve Clark, Xiaodong Xu, and Scott A. Crooker. Probing the influence of dielectric environment on excitons in monolayer WSe2: Insight from high magnetic fields. *Nano Letters*, 16(11):7054–7060, 2016. PMID: 27718588.

[150] H. P. Wagner, A. Schätz, R. Maier, W. Langbein, and J. M. Hvam. Coherent optical nonlinearities and phase relaxation of quasi-three-dimensional and quasi-two-dimensional excitons in zns_xse_{1-x}/ZnSe structures. *Phys. Rev. B*, 56:12581–12588, Nov 1997.

[151] L. Schultheis, A. Honold, J. Kuhl, K. Köhler, and C. W. Tu. Optical dephasing of homogeneously broadened two-dimensional exciton transitions in GaAs quantum wells. *Phys. Rev. B*, 34:9027–9030, Dec 1986.

[152] C. Robert, D. Lagarde, F. Cadiz, G. Wang, B. Lassagne, T. Amand, A. Balocchi, P. Renucci, S. Tongay, B. Urbaszek, and X. Marie. Exciton radiative lifetime in transition metal dichalcogenide monolayers. *Phys. Rev. B*, 93:205423, May 2016.

[153] Lucio Claudio Andreani, Giovanna Panzarini, Alexey V. Kavokin, and Maria R. Vladimirova. Effect of inhomogeneous broadening on optical properties of excitons in quantum wells. *Phys. Rev. B*, 57:4670–4680, Feb 1998.

[154] Voigt J. Influence of spatial dispersion on the transmission spectra of cds single crystals. *physica status solidi (b)*, 64(2):549–556, 1974.

[155] A. Bosacchi, B. Bosacchi, and S. Franchi. Polariton effects in the exciton absorption of gase. *Phys. Rev. Lett.*, 36:1086–1089, May 1976.

[156] Xu Fang, Kevin F. MacDonald, and Nikolay I. Zheludev. Controlling light with light using coherent metadevices: all-optical transistor, summator and invertor. *Light: Science &Amp; Applications*, 4:e292, May 2015. Original Article.

[157] Tomasz Jakubczyk, Valentin Delmonte, Maciej Koperski, Karol Nogajewski, Clment Faugeras, Wolfgang Langbein, Marek Potemski, and Jacek Kasprzak. Radiatively limited dephasing and exciton dynamics in $MoSe_2$ monolayers revealed with four-wave mixing microscopy. *Nano Letters*, 16(9):5333–5339, 2016. PMID: 27517124.

[158] Y. D. Chong, Li Ge, Hui Cao, and A. D. Stone. Coherent perfect absorbers: Time-reversed lasers. *Phys. Rev. Lett.*, 105:053901, Jul 2010.

[159] Jianfa Zhang, Kevin F. MacDonald, and Nikolay I. Zheludev. Controlling light-with-light without nonlinearity. *Light: Science &Amp; Applications*, 1:e18, Jul 2012. Original Article.

[160] Fei He, Kevin F. MacDonald, and Xu Fang. Coherent illumination spectroscopy of nanostructures and thin films on thick substrates. *Opt. Express*, 26(10):12415–12424, May 2018.

[161] Hui Deng, Hartmut Haug, and Yoshihisa Yamamoto. Exciton-polariton bose-einstein condensation. *Rev. Mod. Phys.*, 82:1489–1537, May 2010.

[162] Bo Zhang. *Low Dimensional Polariton Systems in Subwavelength-Grating Based Microcavities*. PhD thesis, University of Michigan, 2015.

[163] Bo Zhang, Zhaorong Wang, Sebastian Brodbeck, Christian Schneider, Martin Kamp, Sven Hfling, and Hui Deng. Zero-dimensional polariton laser in a subwavelength grating-based vertical microcavity. *Light: Science &Amp; Applications*, 3:e135, 2014.

[164] Jonathan Keeling, P. R. Eastham, M. H. Szymanska, and P. B. Littlewood. Polariton condensation with localized excitons and propagating photons. *Physical Review Letters*, 93(22):226403, 2004.

[165] M. H. Szymaska, J. Keeling, and P. B. Littlewood. Nonequilibrium quantum condensation in an incoherently pumped dissipative system. *Physical Review Letters*, 96(23):230602, 2006.

[166] Dimitri Pimenov, Jan von Delft, Leonid Glazman, and Moshe Goldstein. Fermi-edge exciton-polaritons in doped semiconductor microcavities with finite hole mass. *Phys. Rev. B*, 96:155310, Oct 2017.

[167] M. Litinskaia, G.C. La Rocca, and V.M. Agranovich. Inhomogeneous broadening of polaritons in high-quality microcavities and weak localization. *Phys. Rev. B*, 64:165316, Oct 2001.

[168] M. Gurioli, F. Bogani, D. S. Wiersma, Ph. Roussignol, G. Cassabois, G. Khitrova, and H. Gibbs. Experimental study of disorder in a semiconductor microcavity. *Phys. Rev. B*, 64:165309, Oct 2001.

[169] Steven C. Moss, J. Ryan Lindle, H. J. Mackey, and Arthur L. Smirl. Measurement of the diffusion coefficient and recombination effects in germanium by diffraction from opticallyinduced picosecond transient gratings. *Applied Physics Letters*, 39(3):227–229, 1981.

[170] Hui Zhao, Sebastian Moehl, and Heinz Kalt. Coherence length of excitons in a semiconductor quantum well. *Phys. Rev. Lett.*, 89:097401, Aug 2002.

[171] Hui Deng, Glenn S. Solomon, Rudolf Hey, Klaus H. Ploog, and Yoshihisa Yamamoto. Spatial coherence of a polariton condensate. *Phys. Rev. Lett.*, 99:126403, Sep 2007.

[172] S. Miyata. *Nonlinear Optics: Fundamentals, Materials and Devices*. Elsevier Science, 2012.

[173] David Mendlovic and Haldun M. Ozaktas. Fractional fourier transforms and their optical implementation: I. *J. Opt. Soc. Am. A*, 10(9):1875–1881, Sep 1993.

[174] Haldun M. Ozaktas and David Mendlovic. Fractional fourier optics. *J. Opt. Soc. Am. A*, 12(4):743–751, Apr 1995.

[175] C. Candan, M. A. Kutay, and H. M. Ozaktas. The discrete fractional fourier transform. *IEEE Transactions on Signal Processing*, 48(5):1329–1337, May 2000.

[176] A. Laiacano. Fractional fft implementation in python. `https://github.com/alaiacano/frfft`, 2012.

[177] Mark Steger, Gangqiang Liu, Bryan Nelsen, Chitra Gautham, David W. Snoke, Ryan Balili, Loren Pfeiffer, and Ken West. Long-range ballistic motion and coherent flow of long-lifetime polaritons. *Phys. Rev. B*, 88:235314, Dec 2013.

[178] G. Bastard, E. E. Mendez, L. L. Chang, and L. Esaki. Exciton binding energy in quantum wells. *Phys. Rev. B*, 26:1974–1979, Aug 1982.

[179] Michel I. Dyakonov. *Basics of Semiconductor and Spin Physics*, pages 1–28. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[180] Jerome Faist, Federico Capasso, Deborah L. Sivco, Carlo Sirtori, Albert L. Hutchinson, and Alfred Y. Cho. Quantum cascade laser. *Science*, 264(5158):553–556, 1994.

[181] Claire Gmachl, Federico Capasso, Deborah L Sivco, and Alfred Y Cho. Recent progress in quantum cascade lasers and applications. *Reports on Progress in Physics*, 64(11):1533, 2001.

[182] Rohan Singh. *Excitons in Semiconductor Quantum Wells Studied Using Two-Dimensional Coherent Spectroscopy*. PhD thesis, University of Colorado Boulder, 2015.

[183] M. Z. Maialle, E. A. de Andrada e Silva, and L. J. Sham. Exciton spin dynamics in quantum wells. *Phys. Rev. B*, 47:15776–15788, Jun 1993.